Colour constancy in dichromats and trichromats: dependence on task

CHI, MAN,CHUN,ALAN

Colour constancy in dichromats and trichromats: dependence on task

Man Chun Alan Chi

Department of Psychology

Durham University

United Kingdom

A thesis submitted for the degree of

Master of Science (by Thesis)

March 2021

Declaration

The work contained in this thesis was carried out by the author between 2019 and 2020 while a postgraduate student in the Department of Psychology at Durham University. The present thesis has not been submitted to another university for the award of an academic degree in this form. Except where it has been stated otherwise by reference or acknowledgment, the work presented is entirely the author's own.

Statement of Copyright

Acknowledgements

First and foremost, I would like to thank my supervisors Dr. Maria Olkkonen and Dr. Ulrik Beierholm for supporting me throughout this whole research project. Their expertise and guidance made this thesis possible. In particular, their patience and ability to adapt whilst maintaining the high quality of supervision during this COVID-19 pandemic has been exceptional. I thoroughly enjoyed our talks and sharing our ideas. I would also like to thank Dr. Stacey Aston for generously providing us with her optimal colour spaces model, which formed the basis for the observer models in the computerised simulation.

Secondly, I would like to thank my friends for the social support they have provided (albeit all virtually) and in keeping me sane during this pandemic. Without them, I would have been unable to function normally, let alone write a whole Masters' thesis. Thank you for keeping me company (online).

Finally, I would like to thank my parents for financially supporting me throughout this whole endeavour and providing the funding for me to study and research at Durham University.

Abstract

An important topic in the field of colour vision is the impact of colour vision deficiencies on daily life tasks. Investigating the extent to which colour constancy (i.e. the ability to recognise surface colour under different illuminants) is preserved in colour vision-deficient observers can provide us with insight into the nature and function of trichromatic colour vision.

The first chapter of this thesis provides a summary of the very basics of colour vision, colour vision deficiencies, as well as colour constancy. Studies conducted on the colour constancy abilities of colour-vision-deficient observers versus those with normal colour vision are reviewed.

The second chapter presents and reports the aims and methods of the proposed experiment (which could not take place due to the COVID-19 pandemic). This experiment investigated the colour constancy abilities of trichromats versus dichromats using two different colour constancy tasks (2D achromatic adjustment vs. 3D blocks-copying/selection task) and aimed to show how colour constancy depends on observer type as well as task type.

The third chapter comprises of a computerised simulation. This simulation aimed to model the colour constancy of "ideal" observers when presented with various surfaces and illuminants. These observers involve simulated normal trichromats, anomalous trichromats and dichromats. A variety of yellow, blue, green and red illuminant shifts (from neutral daylight) were used, and surface chromaticity and observer types were compared.

Overall, whilst no three-way interaction between illuminant shift, surface chromaticity and observer type were found in the simulation, strong main effects were found. It is suggested that a combination of simulated and experimental research is needed to understand the colour constancy mechanisms underpinning dichromacy and trichromacy at multiple levels (cone-based, cognitive and computational).

Table of Contents

Chapter 1: Introduction

One of the most crucial aspects of visual perception is colour. Our visual experience of colour is what allows us to appreciate the bright blue colour of the sky on a good summer day, or perhaps marvel at the brilliantly saturated colours of flashing neon lights whilst walking along a crowded city street.

On one hand, colour signals are mathematically derived from the optical interactions between the physical properties of objects and illuminations, and how light composed of differing spectral intensities and wavelengths is reflected onto the human eye and retina. On the other hand, there is something subjective and unique to the nature of colour perception that cannot be explained by the reflected light signal itself – colour is part of the visual experience, constructed by the brain. It is fairly easy for one to describe the colour of an object or scene: this apple is red, that banana is yellow, and the beautiful sunset is orange.

For the visual experience of colour to be practical, the visual system should be able to disentangle information about the invariant colour of objects (i.e. that this apple is red) from the illumination and lighting conditions. The sensory signal at the retina changes constantly depending on the illumination incident on the objects we are viewing, even though the physical properties of the object stay the same. The red apple reflects more short-wavelength light under blue skylight during noon on a bright sunny day, whilst it reflects more long-wavelength light under reddish sunlight during sunset. However, the red apple still appears more or less the same "red" colour under different illuminants. The ability to perceive the invariant colour of the red apple under differing light conditions is called colour constancy, which forms one of the main topics of this thesis.

Colour vision deficiency is the second main topic of this thesis. Whilst most people are born with normal colour vision, there are some who either inherit or acquire colour vision deficiencies. The scope of this thesis will be limited to congenital (i.e., inherited) colour vision

deficiencies. Congenital colour vision deficiencies are colloquially referred to as “colour blindness”, and those with such impairments in colour vision will have a different and reduced colour experience. In the most severe cases, those with colour vision deficiencies are unable to discriminate between red and green hues, or in other cases, blue and yellow hues. These deficiencies undoubtedly result in practical disadvantages (e.g. unable to distinguish between red and green traffic light signals). “Colour-blind” people are usually categorised as either dichromats (i.e. having only two functional cone photoreceptor cell types) or anomalous trichromats (i.e. having three functional cone photoreceptor cell types, with one “defective” cone type), whilst those with normal colour vision are normal trichromats. The focus of this thesis will be on red-green anomalous trichromats and red-green dichromats, as well as normal trichromats. Any mention of colour vision deficiencies further in the thesis will exclusively (unless stated otherwise) mean red-green anomalous trichromacy and dichromacy.

It is the goal of this thesis to review and provide insight into how exactly colour constancy may be affected by various red-green colour vision deficiencies. This thesis is divided into four chapters. The first chapter below provides a review of the basics of colour vision, colour vision deficiencies, as well as colour constancy. The methods used to measure colour vision deficiencies and colour constancy are reviewed and explained. Finally, studies conducted on the colour constancy abilities of colour-vision-deficient observers versus those with normal colour vision are reviewed.

The second chapter presents and reports the aims and methods of the proposed experiment (which could not take place due to the COVID-19 pandemic). This experiment investigates the colour constancy abilities of trichromats versus dichromats using two different colour constancy tasks and aims to show how colour constancy depends on observer type as well as task type.

The third chapter comprises of a computerised simulation. This simulation is aimed to model "ideal" observers when presented with various surfaces and illuminants. These observers involve both simulated trichromats and dichromats, using available cone sensitivity functions. Under a shift of neutral daylight illumination to a series of test illuminants, it was found that estimates of colour constancy in the simulation (measured using a simulated error number) was lowest across all observer types when the test illuminants were yellow and green, but highest for the blue and red test illuminants.

In the fourth and final chapter, the findings in this thesis are discussed with regards to the research done so far in this area, and how these findings can be interpreted in the overall context of past and current colour constancy/colour vision deficiency research. Finally, implications for future research are briefly discussed.

1.1 The basics of colour vision

Light (usually known as "visible light") is the portion of the entire electromagnetic spectrum that can be perceived by the human eye (CIE, 2020). The wavelength of visible light usually ranges from 400nm to 700nm (Figure 1.1). Starting from the shortest wavelength to the longest wavelength of light, under common viewing conditions, colours are usually categorised as: violet (~400nm), blue, green, yellow, orange and red (~700nm).

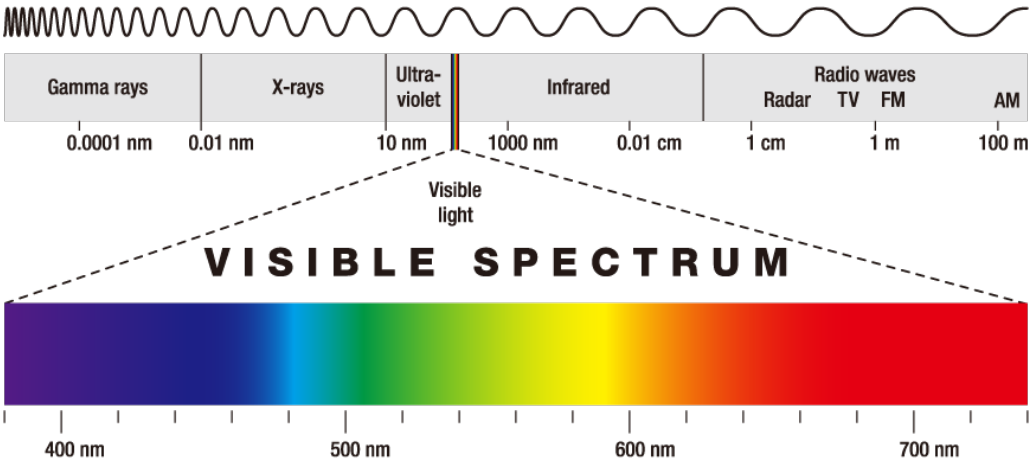

Figure 1.1. The visible portion of the electromagnetic spectrum consists of visible light that can be perceived as colour, from 400nm (purple) to 700nm (red). Source: [Infographic: What is light?]. (2018). https://www.benq.eu/content/dam/b2c/en/knowledge-center/what-is-accurate-color/02-what-is-light.png

The very first step of how colour and light is perceived lies in the physiology of the human eye and retina. As light is reflected off an object, and enters the through the pupil of the eye, this light hits the retinal cells at the back of the eye (Fig 1.2 below).

The transduction of light energy into electrical signals (which then are transformed into action-potentials in the ganglion cells) via the absorption of light in the retinal photoreceptor cells is the first part of the neural process in the perception of colour (Gegenfurtner, 2003).

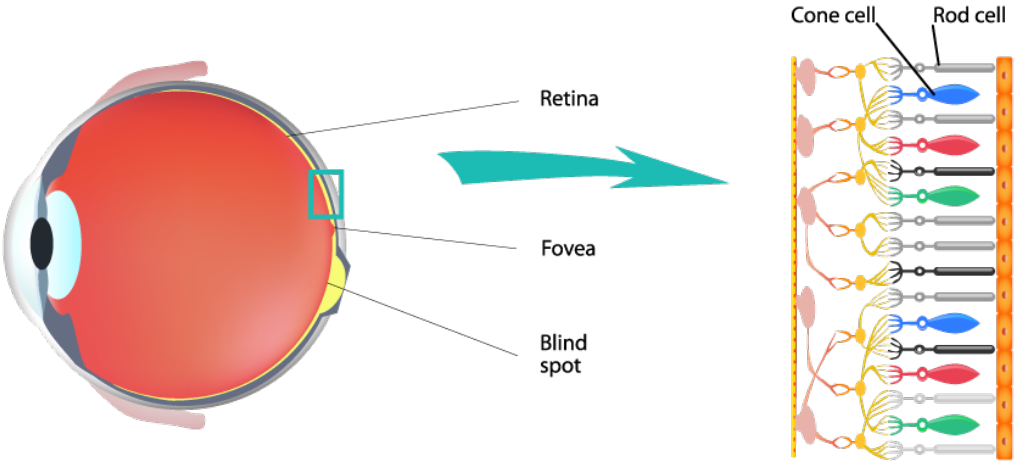

Figure 1.2. The human retina contains four photoreceptor cell types (rods, and three cone types) which are responsible for light perception. Source: [Infographic: What is colour?]. (2018). https://www.benq.eu/content/dam/b2c/en/knowledge-center/what-is-accurate-color/01-what-is-color.png

The photoreceptor cells are usually grouped into two categories: rod and cones. This is because the photoreceptor types have different visual pigments: rods have rhodopsin, whilst cones have opsins. (Kefalov, 2003). Perception of light is mediated via the four photoreceptor cell types that are found in the retina: 1. Rods, 2. S-cones, 3. M-cones, and 4. L-cones. The three cone cell types: S, M, and L-cones, are responsible for detecting short, medium and long-wavelengths of light respectively.

Scotopic vision, which only operates under low light levels ($\sim 1 \times 10^{-6}$ until 1×10^{-2} $\mathrm{cd/m^2}$), is mediated by rod cells. In this state, the human visual system is monochromatic and can only see in shades of grey (Westheimer, 1965).

Photopic vision, on the other hand, only operates under sufficiently high light levels (>1 cd/m^2) and is entirely cone mediated (Willmer, 1961). Visual acuity is high, and colour vision is active in this state. Therefore, this thesis will focus specifically on photopic vision (and subsequently colour perception) mediated by the S, M and L cones. As there are three cone types responsible for colour vision, it is commonly referred to as "trichromatic" colour vision (i.e. tri- for three cone types, -chromatic referring to colour).

As seen from Figure 1.3 above, the spectral sensitivities of individual rods and cones do not cover the entire visible spectrum (measurements by Bowmaker & Dartnall, 1980) – each photoreceptor cell has a peak sensitivity (420, 534 and 564nm for the S, M and L cones respectively, and 498nm for the rods). Whilst these peak sensitivities are separate and distinct from one another, there is a lot of overlap in terms of the overall spectra.

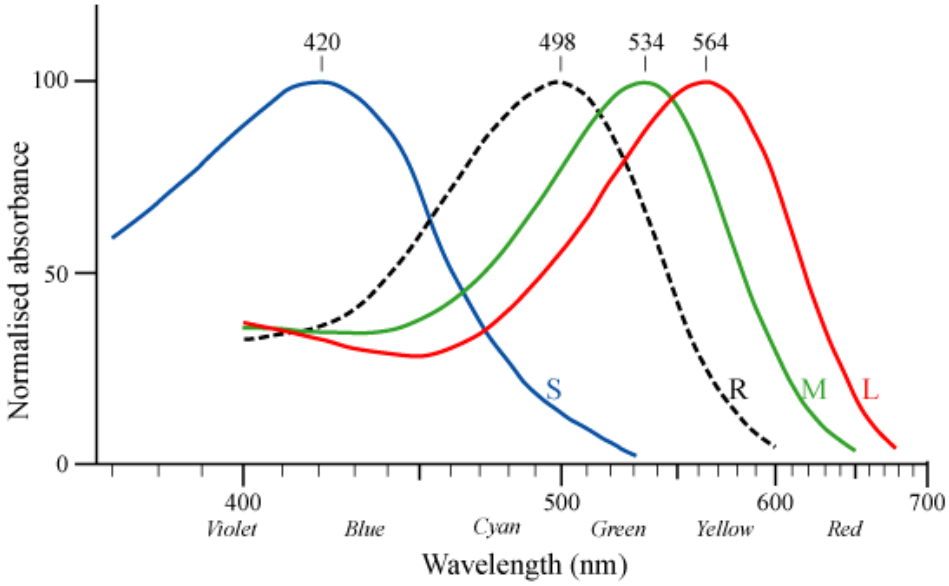

Figure 1.3. Normalised spectral absorbance curves for the S, M and L-wavelength sensitive cones (labelled S, M and L respectively), along with the rods (labelled R), based on measurements from Bowmaker & Dartnall (1980). Source: [Spectral sensitivity curves]. (2009) http://www.academickids.com/encyclopedia/index.php/Image:Cone-response.png

Individual photoreceptor cell types act as photon catches to determine the number of photons that are absorbed by those cells (Smithson, 2005); they can only distinguish the intensity of light and not wavelengths, as wavelength information is lost. Take the case of

scotopic vision, under low light levels, when only the rods are active, resulting in the inability to reliably discriminate between colours. In this state, we can only see in shades of grey. This shows the importance of multiple cone types in colour perception in photopic vision – information from the S, M and L cones are processed via later neural stages in order to extract, amongst other things, wavelength information.

Cone outputs are processed first through the retinal ganglion cells (Berson, 2007), then through the Lateral Geniculate Nucleus (LGN), to the striate cortex (V1). At this point, colour is then processed in conjunction with other visual attributes, such as orientation, depth and motion (Gegenfurtner, 2003), leading to the perceivable and subjective experience of colour.

Whilst a majority of people have normal trichromatic colour vision, exceptions to this at the cone level are detailed in the next section.

1.2 The basics of colour vision deficiency

Colour vision deficiencies (CVD), colloquially referred to as “colour blindness”, occurs when one or more cone photoreceptor cells are defective or non-functional, resulting in impaired colour discrimination and a greatly decreased spectrum of visible colours (Maxwell, 1857). Whilst the majority of people are trichromats (i.e. having normal trichromatic colour vision), a minority of people are dichromats (i.e. only two functional cones) or anomalous trichromats (i.e. two functional cones, one abnormal or “defective” cone). Both dichromats and anomalous trichromats are considered to be colour-vision-deficient (Simunovic, 2010). Whilst there are colour vision deficiencies like achromatopsia (rod monochromacy, a complete lack

of cone cells in the retina) or cone monochromacy (retina only has one type of cone), the focus of thesis will be on dichromacy and anomalous trichromacy.

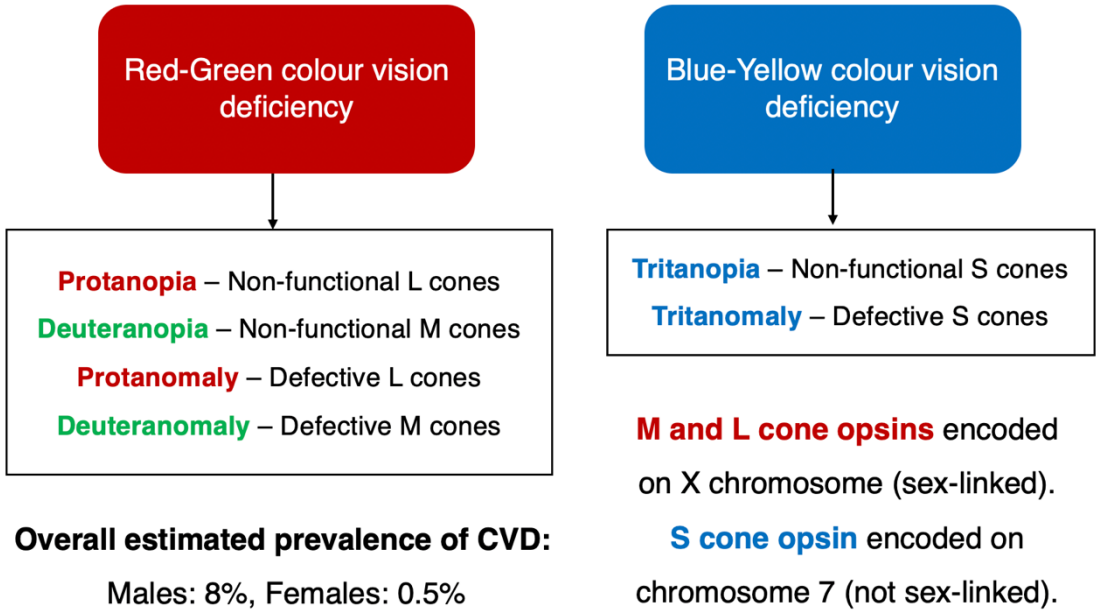

Figure 1.4. Colour vision deficiencies (CVD), or colour blindness, can affect the function of one of three cone types. The level of CVD severity depends on whether the cone is non-functional (protanopia, deuteranopia or tritanopia), or simply just defective (protanomaly, deuteranomaly, tritanomaly).

Figure 1.4 above shows that colour vision deficiencies can be categorised into: 1. Red-Green or 2. Blue-Yellow. The colour names refer to the spectrum of colours that are harder to discriminate for the observer. The severity of the CVD depends on how cone function is affected.

From a genetics standpoint, as M and L cone opsins are coded on the X chromosome, red-green CVD is sex-linked and affects men disproportionally. On the other hand, as S cone opsins are encoded on chromosome 7 are not sex-linked, men and women are affected equally by blue-yellow CVD. However, in terms of overall numbers, blue-yellow CVD is extremely rare compared to Red-Green CVD, with a prevalence of 0.03% compared to 6% (Crone, 1968;

Deeb, 2004; Neitz, 2011). This is the main reason many studies have only included red-green colour-vision-deficient observers.

Protanopia, Deuteranopia and Tritanopia (Figure 1.4) are all considered "colour blindness", as the L, M and S cones are completely non-functional in their respective cases (for the cone sensitivity curves see Figure 1.5 below). People with protanopia, deuteranopia and tritanopia are all considered dichromats. Dichromats perform worse than trichromats on colour discrimination tasks (Álvaro et al., 2015) and may find it impossible distinguishing between the colours on the red-green or blue-yellow spectrum.

Protanomaly, Deuteranomaly and Tritanomaly are considered a milder form of CVD, as the L, M or S cones are "defective" but still functional, in other words, abnormal. This abnormality usually manifests itself as the affected cone type having an altered spectral sensitivity, usually shifted closer to the other cone types (Pokorny & Smith, 1977; DeMarco, Pokorny & Smith, 1992) – see top middle and top right graphs in Figure 1.5 below. This results

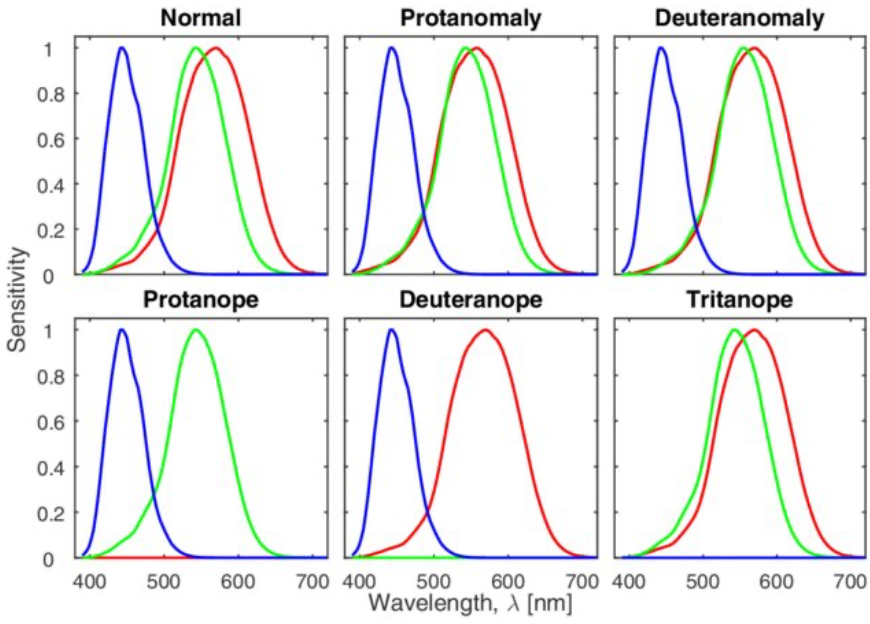

Figure 1.5. Normalised cone sensitivities (from top left to bottom right) for 1. Trichromat, 2. Protanomalous trichromat, 3. Deuteranomalous trichromat, 4. Protanope, 5. Deuteranope, 6. Tritanope. Colours for the curves represent responses for the S (blue), M (green) and L (red) cones respectively. Source: Shrestha (2016).

in a reduction of colour space and discrimination for the anomalous trichromat compared to normal trichromats, albeit the reduction in colour discrimination not as severe as dichromats.

Due to blue-yellow colour vision deficiencies being extremely rare (Crone, 1968), the focus of this thesis will be on red-green colour vision deficiencies and will specifically address the colour constancy abilities (this will be explained in section 1.4) of five observer types: 1. normal trichromats, red-green anomalous trichromats (2. protanomalous and 3. deuteranomalous trichromats) as well as red-green dichromats (4. protanopes and 5. deuteranopes).

1.3 Methods of measuring colour vision deficiencies

Whilst there are a large number of clinical colour vision tests that are available for measuring colour vision deficiencies (Dain, 2004), this section will focus on explaining the following three common methods, each of which utilise a different testing paradigm (in brackets):

1. The Ishihara test (Pseudo-isochromatic plates test)
2. Farnsworth-Munsell 100-hue Test (Arrangement test)
3. Anomaloscope (Colour matching test)

First is the Ishihara test (Ishihara, 1917). This test is frequently referred to as the "gold standard" for rapidly identifying and screening for red-green colour vision deficiencies (particularly those of a congenital nature – i.e. inherited) due to its ease of use and efficiency (Birch, 1997). The Ishihara test is the most well-known and most commonly used colour vision test worldwide (Dain, 2004).

The full test consists of 38 printed plates, each of which are construed in such a way that the coloured dots which form the main number (usually referred to as the "object" which

the observer has to correctly identify) is surrounded by a coloured background of the same luminance (to avoid non-colour related cues – this is what is meant by pseudo-isochromatic).

The observer's task in the Ishihara test is to correctly read off the number on each plate. Ideally, each plate should be presented under daylight illumination, and to only give three seconds per plate for an answer (Ishihara, 1972). The chromaticities of the dots and background are chosen such that those with red-green colour vision deficiencies will have a difficult time identifying the number, and answer a different number compared to trichromats. See Figure 1.6 below.

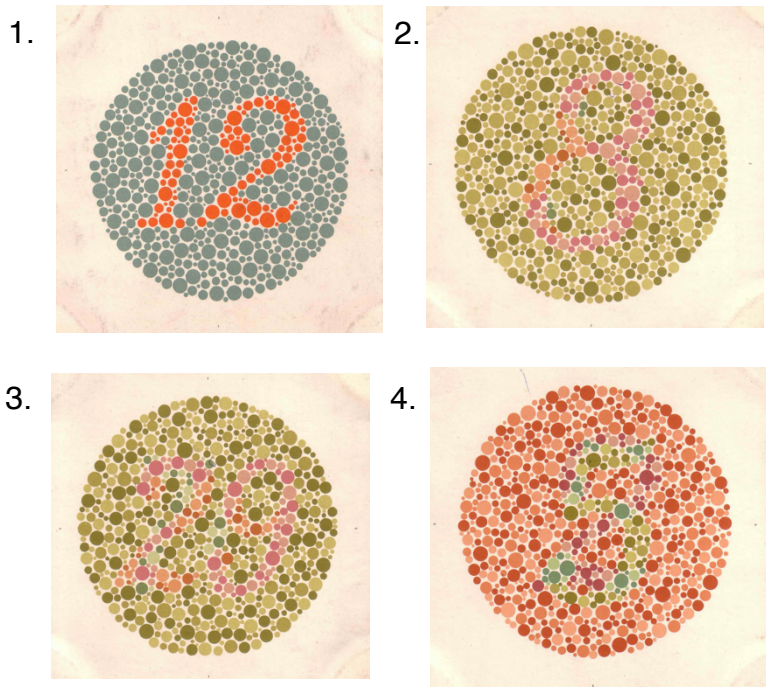

Figure 1.6. The first four plates from the Ishihara test, labelled 1 to 4. Those with normal trichromatic colour vision should see the numbers 12, 8, 29 and 5 respectively, whereas those with red-green deficiencies might see 12, 3, 70 and 2 respectively. The plates here are only an approximation of what the observer might see. The real test plates should be presented under a standard daylight illumination. Source: Ishihara (1972).

Whilst the Ishihara test is a quick and reliable diagnostic measure of whether an individual has a red-green colour vision deficiency, it is unable to accurately measure the severity of the deficiency, as well as not being to detect minimally anomalous trichromats (Dain,

2004). Thus, initial CVD screening is usually done by the Ishihara test, but to make an accurate severity diagnosis, methods such as the Farnsworth-Munsell 100-hue test, as well as the anomaloscope are used.

The second method detailed below is that of the Farnsworth-Munsell 100-hue test. The Farnsworth-Munsell 100-hue test is classified as an arrangement test, which involves the observer to sort colours into a sequence based on the hues. The colours in this test are based on the Munsell system (Munsell, 1912), which separates colour attributes into perceptually uniform dimensions: hue, chroma and lightness. The Farnsworth-Munsell 100-hue test (Farnsworth, 1943) only utilises colours that vary in hue, whilst keeping the other Munsell dimensions the same. Even though there are 100 hues in the Munsell system, difficulties in discriminating between certain hues eventually led Farnsworth to remove 15 hues in order to make the difficulty of colour discrimination more uniform. The final Farnsworth-Munsell 100-hue test, therefore, has 85 hues (even though its official name is the 100-hue test).

Figure 1.7. The physical version of the Farnsworth-Munsell 100-hue test. 85 hue caps are divided into four rows. The hues on either end of each row are fixed, and the observer has to rearrange the caps such that the hues are perceived to have a gradual progression. Source: [Image of the Farnsworth-Munsell 100-hue test]. (n.d.). https://www.bernell.com/product/LV3553000/94

These 85 hues are then divided into four rows/boxes (Figure 1.7). Each row contains a fixed hue at the end. The observer is asked to rearrange the moveable caps in accordance with the hues, such that the hues are perceived to have a gradual progression from the start to end of each row.

To determine the severity and type of CVD, the order in which the hues were arranged by the observer are analysed. Results (i.e. the error scores quantifying the mis-orderings) are plotted on a polar plot, from then which the confusion axis is determined – this determines the type of CVD. In addition, the greater the intensity and frequency of the deviation from the center, the higher the severity of CVD. Below in Figure 1.8 are example results of the Farnsworth-Munsell 100-hue test for a trichromat, protanope and deuteranope respectively.

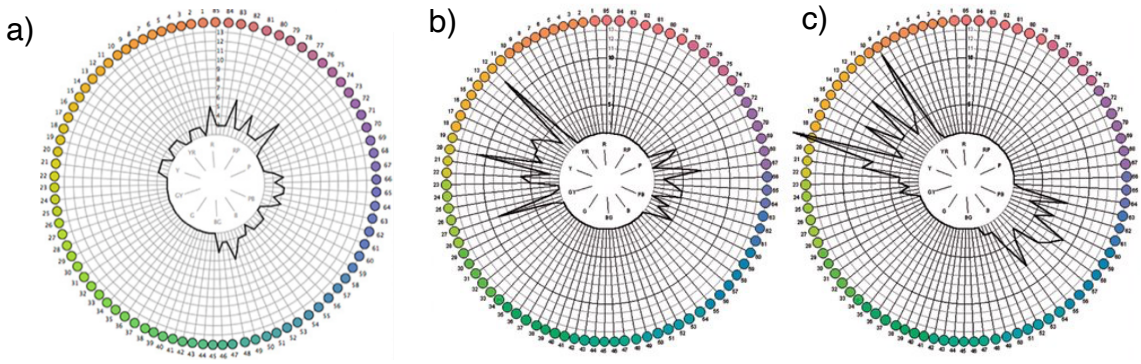

Figure 1.8. Results for the Farnsworth-Munsell 100-hue test for a) normal trichromat, b) protanope and c) deuteranope. The error scores quantifying the mis-orderings are plotted on the polar plot. An observer with perfect trichromatic colour vision and colour discrimination would have a perfect circle plotted in the middle with no deviations. Source: Ferreira (2017).

Whilst the Farnsworth-Munsell 100-hue test is very precise when it comes to determining slight deviations in colour discrimination, and is a useful colour vision assessment, because of the difficulty of the test itself, it sometimes fails to distinguish between normal trichromats (with poorer colour discrimination) and mild anomalous trichromats (with better colour discrimination), and may not differentiate between protanopes or deuteranopes if their confusion axes are too similar (Birch, 1989; Lagerlöf, 1991). Arrangement tests like the Farnsworth-Munsell 100-hue test should be combined with pseudo-isochromatic tests (e.g. Ishihara test) and colour matching tests (e.g. anomaloscope) for a more accurate diagnosis.

The third method for measuring colour vision deficiencies is the anomaloscope. One of the most common is Nagel's anomaloscope. This instrument was invented by Willibald Nagel

(Nagel, 1907) and aims to quantify the level of CVD using the Rayleigh matching equation (Rayleigh, 1881).

The Rayleigh equation (Rayleigh, 1881) can be expressed simply as: Red + Green = Yellow. Essentially, spectrally pure (single wavelength) lights can be additively mixed together

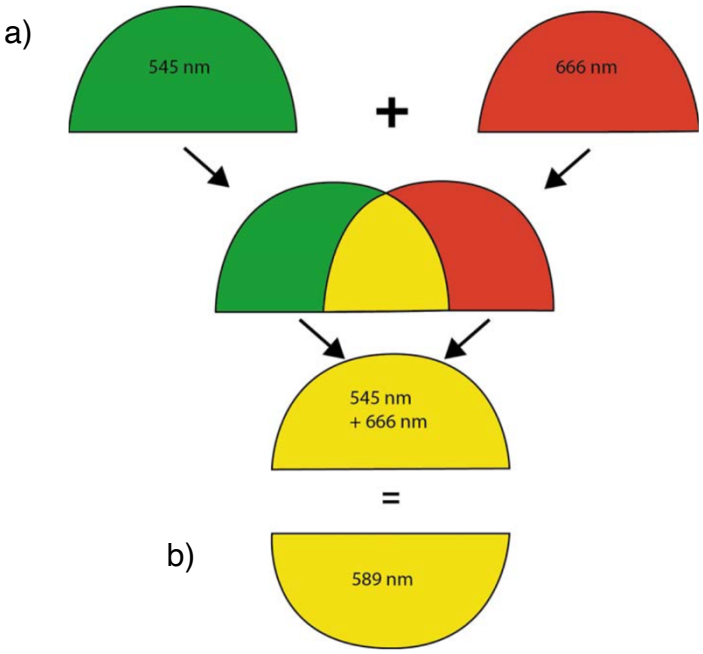

Figure 1.9. Rayleigh equation for colour matching: Green (545nm) + Red (666nm) = Yellow (589nm). The top part (a) shows the green and red lights additively mixing together to match the spectrally pure yellow in the bottom (b). The mixture of red and green can be varied, or, the brightness of the yellow can be adjusted to perform a colour match. Source: Krastel (2015).

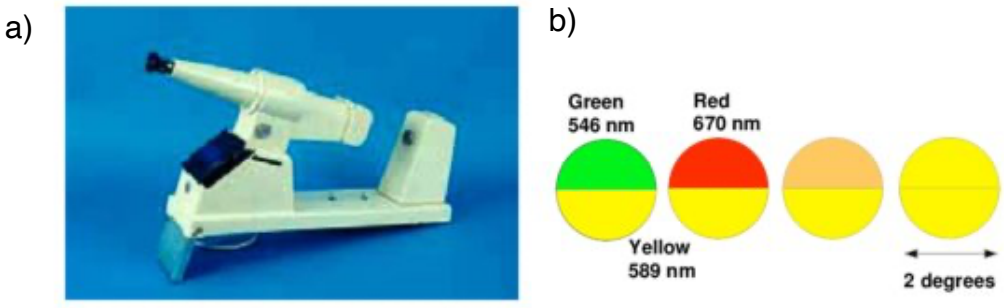

Figure 1.10. The Schmidt and Haensch Model I Nagel anomaloscope (a). The observer has to look through the eye piece and control two stimulus knobs (b) which adjust the ratio of: 1. The brightness of yellow light (589nm) required to match a given additive mixture of red and green lights, and 2. The amount of red (670nm) and green (546nm) light for the additive mixture. Source: Barbur et al. (2009).

to obtain a "colour match" (Figure 1.9) – this colour matching forms the basis of the anomaloscope test.

Figure 1.10 above shows the first commercially available anomaloscope, the Schmidt and Haensch Model I Nagel anomaloscope (Barbur et al., 2009). The observer's task is to look through the eye piece and match the colours using two stimulus knobs which adjust the ratio of the red-green light mixture (top part of the circle) or the brightness of the yellow light (bottom part). Within the light mixture field (top), the relative amounts of red and green lights needed to spectrally match the spectrally pure yellow light (bottom) is dependent on the spectral curves of the observers' L and M cones. As such, observers with CVD will either require more red or green light in the mixture field in order to match the colour of the yellow field, depending on how much their L and M cones deviate from the normal trichromat spectral curves (Barbur & Rodriguez-Carmona, 2017).

The anomaloscope permits for the full examination of anomalous trichromacy and dichromacy using this colour matching paradigm (based on the Rayleigh equation) and can only be administered by a professional who has undergone extensive training. Due to the number of matches that have to be made for a full examination, anomaloscope testing is much more time-consuming than the Ishihara test or Farnsworth-Munsell 100-hue test. However, when administered by a trained and skilled examiner, compared to all other methods, the anomaloscope is the gold-standard clinical instrument in diagnosing and classifying colour vision deficiencies (National Research Council, 1981).

In sum, the three methods used described here (Ishihara test, Farnsworth-Munsell 100-hue test, and the anomaloscope) are all used in different settings to screen, identify, and measure CVDs. There is no one test which can fulfil all needs. Each method has its strengths and weaknesses and should be used in conjunction with one another depending on the demands of the situation (National Research Council, 1981; Dain, 2004).

1.4 The basics of colour constancy

Colour constancy is the ability of the human visual system to perceive invariant material colour despite changes in the intensity and spectral composition of the illumination (Foster, 2011). A critical part of colour vision, colour constancy is essential in visual object recognition. Colour constancy is responsible for the fact that we know the pelargonium plant placed under reddish sunlight in the image on the left (Figure 1.11 below) is the same material colour as the plant on the right (under bluish skylight), despite the drastically different reflected colour signals (as shown by the spectral radiance graphs below each image representing the spectral intensity and composition of light reflected by the petals denoted by the white arrows within each image).

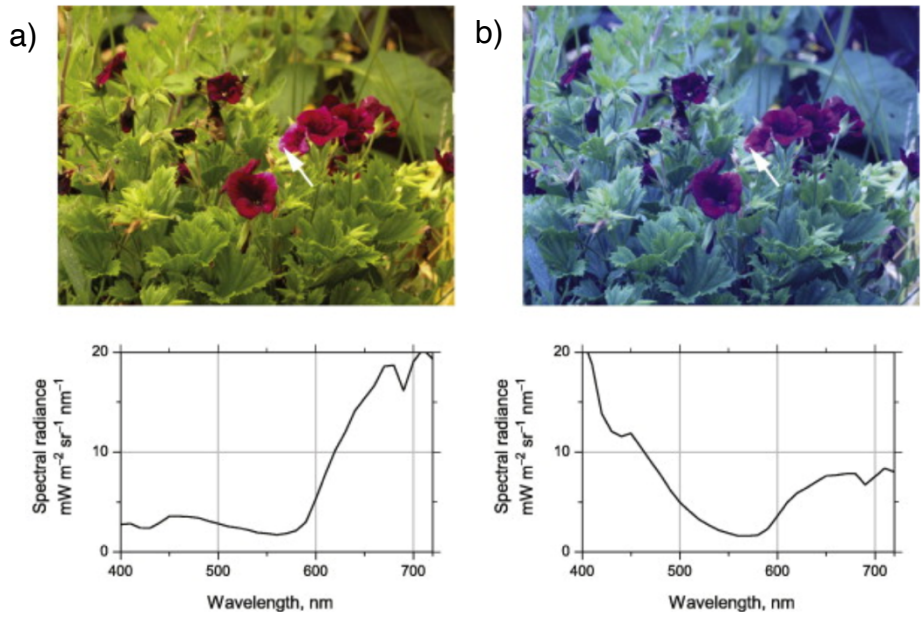

Figure 1.11. Colour constancy under two drastically different illuminations. In the image on the left (a), a pelargonium is illuminated by reddish direct sunlight, and the image on the right (b) shows the same flower under a bluish skylight illuminant. The observer is able to infer the material colour of the flower. Source: Foster (2011).

Whilst it might seem an intuitive task for the visual system to recognise material colour, colour constancy is a mathematically ill-posed problem. This is because for any given sensory signal, there can be an infinite number of illumination and surface combinations. Figure 1.12 below illustrates the problem: for any given constant physical property of an object or surface, the reflected light signal that is incident on the retina is a product of the surface reflectance $R(\lambda)$, and the illumination $E(\lambda)$.

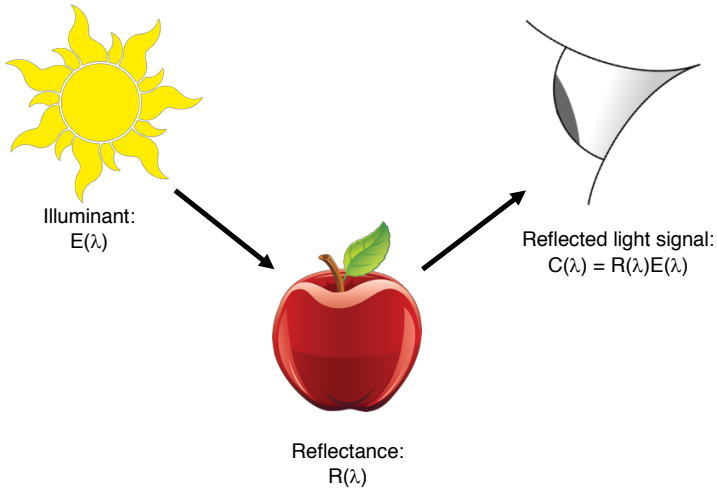

Figure 1.12. For a constant surface or object, the reflected light signal (which contains colour information) depends on both the reflectance and illuminant. The reflected light signal can be drastically different depending on the illuminant the object is presented under.

Despite the different reflected colour signals, the human visual system is somehow able to solve this ill-posed problem by accurately estimating the surface colour/reflectance most of the time, despite with limited available information (resultant cone absorptions), rather than possessing full information about the spectral composition of the illuminant and surface.

In colour vision, this interaction between the physical aspects of the surface and illuminant and the biological aspects of the observer is known as the "colour conversion" process (Helson, 1938), whereby the resultant cone absorptions change based on changes in

the reflected light signal $C(\lambda)$, which changes based on the spectral composition and intensity of the illumination $E(\lambda)$.

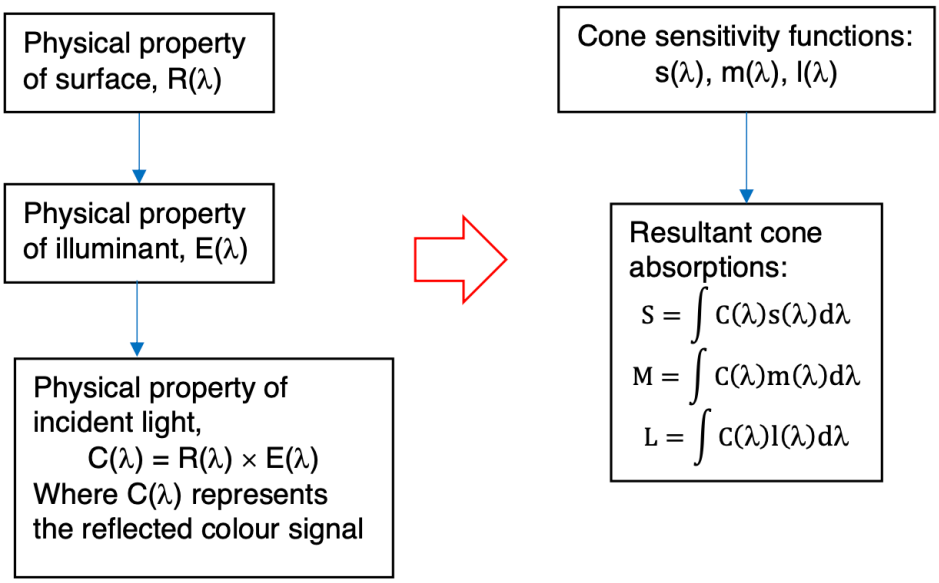

Figure 1.13. The human visual system only has available information from the resultant cone absorptions but is somehow able to accurately estimate the physical properties of surfaces and illuminants (in this case, the surface reflectance and its associated material colour).

To sum up the problem of colour constancy: the human visual system only has access to the resultant cone absorptions, essentially three numbers (See Fig 1.13 above). To "convert" these cone absorptions (S, M, L) into a proper representation and visual experience of surface colour, and to separate that from the illumination, colour constancy mechanisms are required (Smithson, 2005).

Investigating colour constancy mechanisms can provide insight into how the sensory inputs (cone absorptions) give rise to the perception of invariant material colour. At the retinal level, sensory mechanisms such as physiological adaptation were proposed to be essential in colour constancy. For example, the famous von Kries coefficient law (von Kries, 1878) suggested that colour constancy might be explained via a simple linear model, where the same object viewed under different illuminants simply results in a multiplicative change in the cone

absorptions independently for each cone. This idea of receptor scaling is still used in current opponent-process models of the early stages of colour vision (Pugh & Mollon, 1979).

However, it is clear that physiological adaptation in the retina is insufficient for explaining the entire scope of colour constancy. Whilst lesion and neuroimaging studies have tried to locate a "functional colour constancy" area in the human brain (Land et al., 1983; Ruttiger et al., 1999), other researchers like Wandell et al. (1999) suggest that instead of looking for a specialised "colour constancy" area, it might be more useful to employ neuroimaging techniques to track the processing of visual information along specialised visual pathways.

Recent studies suggest that colour constancy mechanisms can potentially be driven by a combination of specific extrastriate neural mechanisms that construct a stable representation of material colour, that are more nuanced and complex than the simple colour-selective cells found in the striate cortex (Johnson et al., 2008). These neural mechanisms would account for a large proportion in computing colour constancy, with a small proportion being accounted for by sensory mechanisms such as chromatic adaptation (Rinner & Gegenfurtner, 2000; Werner, 2014).

Classically, Land's Retinex theory (Land & McCann, 1971) proposed that the processing of lightness and colour signals involve global interactions between the cortex and the retina. Land's idea that the retina and cortex interact together to produce stable colour perception has been to some extent verified by modern neuroimaging studies, which have shown that monocularly-driven neural mechanisms involving the V1, V2, V3 and V4/V4A areas of the cortex are also activated in the computation of colour constancy (Barbur & Sprang, 2008; Bannert & Bartels, 2017), allowing the observer to rapidly detect a change in colour that is due to an illuminant shift, rather than a change in the surface reflectance.

These neural mechanisms have to work in conjunction with the more basic properties of the retina. After all, the reflected colour signals have to reach the retinal ganglion cells before

any processing by the cortex can begin (Zeki & Marini, 1998). As soon as the light reaches the retina, however, it appears that colour constancy mechanisms can operate at an extremely rapid and potentially nonconscious level (Foster et al., 1992; Norman et al., 2014).

However, whilst studies have shown that extrastriate areas of the brain are activated during colour constancy (Conway, 2001; Barbur & Sprang, 2008; Bannert & Bartels, 2017), the nature of the interactions between retinal, striate and extrastriate cortical mechanisms and how they each contribute to colour constancy is not yet precisely known. Whilst it is likely that there are differences in cortical colour representations in observers with CVD, the colour constancy tasks described and used in this thesis cannot differentiate between these levels of processing. The focus on this thesis is on how different retinal cone mosaics/phenotypes (i.e. trichromacy, anomalous trichromacy and dichromacy) may affect colour constancy.

1.5 Methods of measuring colour constancy

Currently, there are many different psychophysical methods with which to measure colour constancy. This section will briefly cover and explain the following three methods:

1. Asymmetric colour matching (adjustment task)
2. Achromatic adjustment (adjustment task)
3. Blocks-copying task (object identification task)

First of all are the adjustment tasks: 1. Asymmetric colour matching, and 2. Achromatic adjustment. These are both popular and traditional measures of colour constancy. Whilst both are adjustment tasks (i.e. involve some form of colour matching), asymmetric colour matching focuses on matching to a standard stimulus (explicit measure), whereas achromatic adjustment focuses on matching to an internal standard (implicit measure).

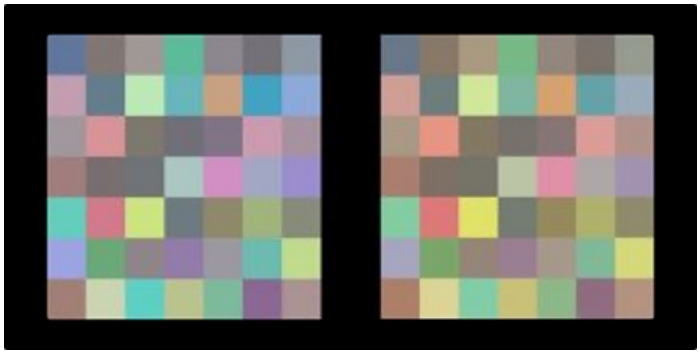

Figure 1.14. Simultaneous asymmetric matching: the same set of square patches (i.e. same surface reflectances), illuminated under two different illuminants. Observers are presented both sets at the same time and asked to adjust and match the central patches to "make it look as if it were cut from the same piece of paper". Source: Reeves (2018).

The first type of traditional method, asymmetric colour matching, is a task whereby an observer is asked to match a sample surface colour under one illuminant, to a test surface colour under a second illuminant (Fig 1.14 above). Asymmetric colour matching can be successive, with the sample and test colours presented one after the other (Brainard & Wandell, 1992), or simultaneous, with the sample and test colours presented together, side-by-side (Arend & Reeves, 1986; Reeves, 2018).

One of the first studies to investigate colour constancy using this matching method was Arend & Reeves (1986). Arend & Reeves (1986) asked the observer to match the test surface colour by making it "look as if it were cut from the same piece of paper [as the sample surface colour]". Arend & Reeves (1986) also asked the observers to match the "hue and saturation" of the sample and test colours. Whilst both the paper and hue-saturation matching tasks were interpreted as matching the subjective experience of colour, the paper matches consistently yielded much higher constancy indices than hue-saturation matches, suggesting that the explicit instructions for a "paper match" tap into colour constancy mechanisms much better than "hue-saturation" matches (Arend & Reeves, 1986; Arend et al., 1991; Smithson, 2005; Radonjić & Brainard, 2016).

On the other hand, achromatic adjustment (Figure 1.15 below) is another traditional method for measuring colour constancy (Rüttiger et al., 2001; Foster, 2003). This type of task involves asking the observer to adjust the colour of a central patch until it appears white or neutral gray (given that the patterns are illuminated differently). This then gives an estimate of how well the observer has perceived invariant surface colour.

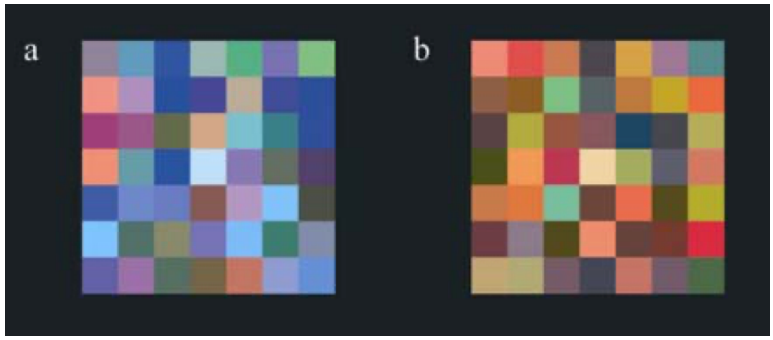

Figure 1.15. An achromatic adjustment task involves asking the observer to adjust the central patch until it appears white. Pattern A is illuminated by blue skylight while pattern B is illuminated by red sunlight. The two central patches show a successful adjustment to white. Source: Foster (2003).

Using achromatic adjustment, chromatic adaptation and colour context effects can be investigated to determine colour constancy. Morgan et al. (2000) showed that the accuracy of implicit measures such as achromatic adjustment can be just as good as explicit measures such as asymmetric colour matching.

In both asymmetric colour matching and achromatic adjustment, colour constancy is typically estimated via an index. Similar to how other perceptual constancies are calculated via the Brunswik ratio (Hurvich & Jameson, 1966), the colour constancy index is calculated with the same principle in mind. The Brunswik ratio can be expressed simply as:

$$Perceptual\ constancy\ index = \frac{R - S}{A - S}$$

Where:

- R is the physical magnitude/intensity of the stimulus chosen as a match

- A is the physical magnitude/intensity of the stimulus chosen under 100% constancy
- S is the physical magnitude/intensity of the stimulus chosen under 0% constancy.

Thus, the ratio becomes zero when the chosen match stimulus R is equal to S, and the ratio becomes 1 when R is equal to A.

In the case of colour constancy, the physical magnitude/intensity is the chromaticity of the stimuli, which can for example be expressed in CIE_{xy} colour space (CIE, 1932). Each "colour" has a specified CIE_{xy} coordinate – each coordinate representing the unique reflected light signal of the stimulus (the product of surface reflectance and illumination). An example how the colour constancy index is calculated is shown in Figure 1.16 below (Granzier & Gegenfurtner, 2012).

From Figure 1.16 below, the colour constancy index can be calculated by the formula:

$$Colour\ constancy\ index = \frac{|d| \times |c| - |b| \times |a|}{|d|^2 + |b|^2}$$

Which gives a number between 0 and 1 (like the Brunswik ratio). This number is then converted into a percentage, and represents how colour constant the observer is under an illuminant shift. This calculation of the index applies to both asymmetric colour matching and achromatic adjustment methods: in the case of asymmetric colour matching, the observer match (actual match – open square) would measure how close the observer has adjusted the central patch of test surface colour to the sample surface colour, whereas in achromatic adjustment,

the observer match would measure how close the observer has adjusted the central patch to a neutral white.

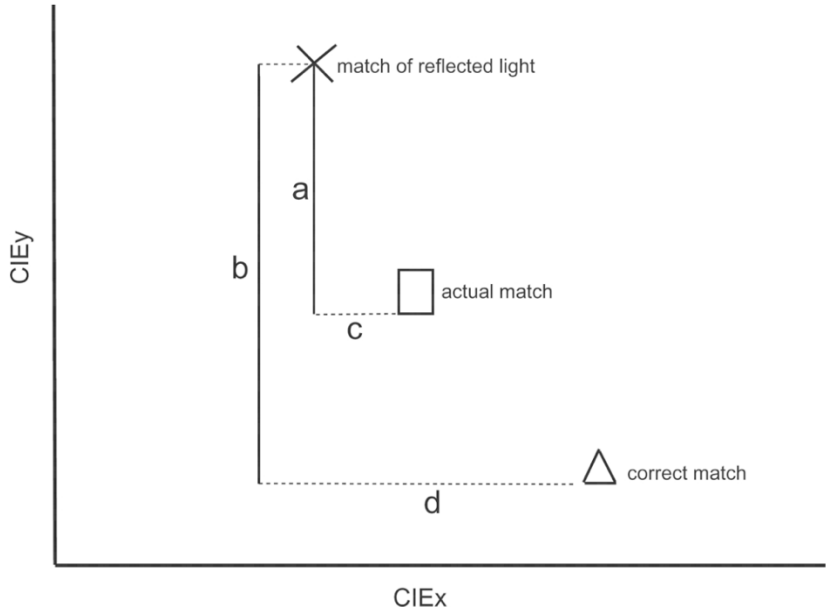

Figure 1.16. Colour constancy index is calculated using the vector values a, b, c and d. a and b are the vertical vector components, representing the change in chromaticity of the CIE y coordinate. c and d are the horizontal vector components, representing the change in chromaticity of the CIE x coordinate. Chromaticities are expressed in CIE (x,y) coordinates. The cross shows the chromaticity of the match of reflected light (i.e. zero constancy, or tristimulus match under an illuminant shift). The open square shows the chromaticity of the actual match made by the observer. The open triangle shows the chromaticity of the correct match (i.e. perfect constancy, perfect match of surface reflectance under illuminant shift). The observer's match (open square) is usually somewhere in between 0 and 100% constancy. Source: Granzier & Gegenfurtner (2012).

All methods that measure colour constancy, including asymmetric colour matching and achromatic adjustment, must have three essential pieces of information, which are required to calculate the colour constancy index. These are the chromaticities of the:

1. Match of reflected light (tristimulus match – 0% constancy)
2. Actual match (observer match)
3. Correct match (perfect surface reflectance match under illuminant shift – 100% constancy)

Whilst achromatic adjustment and asymmetric matching tasks are both traditional ways of measuring colour constancy, they are not without drawbacks. Quantifying colour constancy

via such “unnatural” tasks does not take into how surface colour is utilised in real life. Colour is usually a fixed physical property of the surface – meaning that it is highly unlikely in a real life situation that the colour of an object can simply be manipulated or changed by the observer. Achromatic adjustment and asymmetric matching tasks both use a similar experimental paradigm where the observer has to manually adjust the colour of the object/surface until they feel like it matches under the illumination shift.

To accurately quantify and measure colour constancy, the function of colour constancy and how it operates in real life must be taken into account. As one important function of colour constancy is object selection and identification, the perception of invariant surface colour allows one to accurately select, for example, the ripest bananas, whether it be bananas illuminated under natural blue skylight in an outdoor market, or bananas illuminated under the bright white LEDs in an indoor supermarket. To address this, object identification tasks using forced-choice matching paradigms were developed (Bramwell & Hulbert, 1996; Zaidi & Bostic, 2008), where the observer is shown a fixed set of stimuli and forced to choose the “odd one” out, or state whether the object is the same or different. In the case of Zaidi and Bostic (2008), they used a four-alternative forced-choice task whereby observers were presented four similarly shaped real objects under different illuminants (with three of the four objects having the same surface reflectance/material) and asked to identify and pick the odd object. This object identification task and forced-choice matching paradigm used to measure “functional” colour constancy forms the basis for the blocks-copying task (Radonjić, Cottaris & Brainard, 2015).

An object selection task, nicknamed the “blocks-copying” task was developed by Radonjić, Cottaris and Brainard (2015) to measure functional colour constancy. The blocks-copying task is naturalistic, goal-directed task that mirrors the way in which colour information is used in real life (Fig 1.17 below). Three simulated areas (from left to right): the model, workspace and source are shown to the observer. The model contains four blocks (these are the

sample surfaces), and the workspace contains four placeholders. The observer's task is to select, from amongst the eight blocks in the source, four blocks which match the surface colours of the blocks in the model, and replace the placeholders in the workspace with the four blocks chosen from the source.

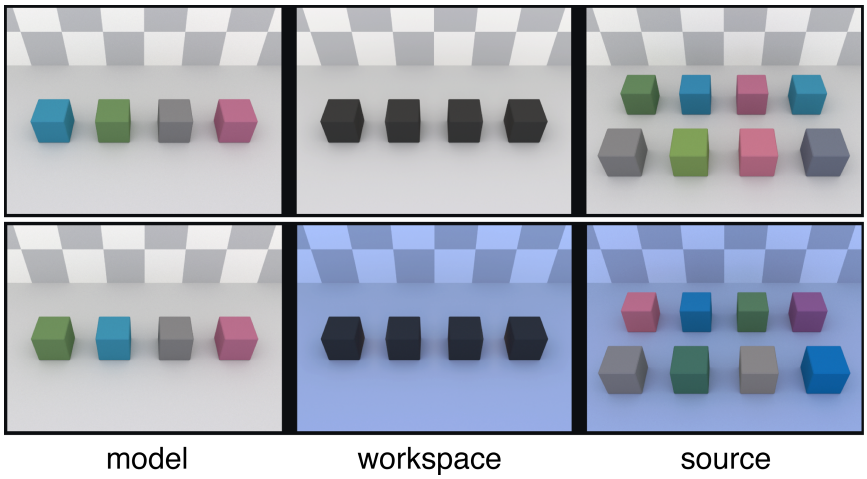

Figure 1.17. Three rendered rooms with simple block-like objects are presented to the participant. Subjects are asked to recreate the arrangement of blocks in the model by replacing the 4 blocks in the workspace from a selection of 8 blocks from the source. The top row shows the same illuminant for all three rooms, whilst the bottom row shows an illuminant change for the workspace and source. Source: Radonjić, Cottaris & Brainard (2015).

The surface colours of all the blocks were determined beforehand for each test and illuminant condition, and the spatial arrangement of the blocks were randomised. Performance on the blocks-copying task (Radonjić, Cottaris & Brainard, 2015) have shown to be comparable to performance on more traditional tasks such as asymmetric colour matching (Foster, 2011), with mean colour constancy indices of 46% and 52% respectively.

In summary, the three methods explained in this section (asymmetric colour matching, achromatic adjustment and blocks-copying task) are all useful and comparable in measuring colour constancy. The achromatic adjustment and blocks-copying task are of particular interest, as these will be used in the experiment described in Chapter 2 of this thesis.

1.6 Colour constancy and colour vision deficiencies

This section will review a number of colour constancy studies (for a quick summary, see Table 1.1 at the end of this section) and provide some discussion points on how colour constancy may, or may not differ in trichromats versus dichromats. The focus will specifically be on normal trichromats, red-green anomalous trichromats and red-green dichromats. Unless specified otherwise, the term "anomalous trichromat" will refer to those with protanomaly and deuteranomaly, and "dichromat" will refer to those with protanopia and deuteranopia.

An important topic in the field of colour vision is the impact of colour vision deficiencies on daily life tasks. Investigating the extent to which colour constancy (i.e. the ability to recognise surface colour under different illuminants) is preserved in colour vision-deficient observers can provide us with insight into the nature and function of trichromatic colour vision. Questions such as the following can be asked:

- Is there any difference in colour constancy performance between normal trichromats, anomalous trichromats and dichromats?
- What potential explanations are there to account for these differences?

Whilst many previous studies have tested colour constancy on red-green dichromats and compared them to normal trichromatic observers, these studies have been inconclusive as to the extent and nature of colour constancy in colour vision deficient observers (e.g. Rüttiger et al., 2001; Baraas et al., 2010; Ma et al., 2016; Alvaro et al., 2017).

It is known that colour constancy is not mediated purely by retinal mechanisms, and that the visual cortex plays a large role in computing surface colour (Rüttiger et al., 1999; Barbur & Sprang, 2008). In the most severe cases of complete colour blindness, such as acquired achromatopsia in Patient BL (Morland et al., 1997), the photoreceptoral and post-receptoral opponent mechanisms are intact, but the pre-striate cortex is damaged. Patient BL's colour constancy performance was abnormally poor for his colour discrimination loss,

indicating that his cortical lesion accounted for a large role in processing colour constancy. On the other hand, the photoreceptoral mechanisms in red-green deficient observers are different than that of normal trichromats, but unlike those with acquired achromatopsia, their striate cortex is intact. It is rather difficult to ascertain how colour constancy mechanisms are impaired in red-green deficient observers, and most explanations usually associate poorer colour discrimination with poor colour constancy.

It is known that colour naming and colour discrimination are typically correlated: the better the discrimination, the better/faster the naming (Montag, 1994; Morland et al., 1997). Colour naming can aid in understanding the effects of colour vision deficiencies on colour constancy. Montag (1994) conducted an experiment on trichromats versus red-green dichromats using a monolexemic (single-word) colour naming task. It was found that shorter stimuli presentation time impaired the dichromats' colour naming abilities. Even under high light levels (to rule out scotopic, rod-based vision), there was still an effect of short duration presentation on colour naming for dichromats, whereas for normal trichromats, no such duration or light level effects were found. It is clear that dichromats perceive less of the colour space than normal trichromats, and this has a significant impact on their colour naming abilities.

The relationship between colour naming, discrimination and constancy was investigated in Morland et al. (1997). This was one of the earliest studies to specifically investigate and compare the colour constancy abilities of normal trichromats, anomalous trichromats and dichromats. Using a colour naming task with five different illuminant shifts to measure colour constancy, they found that for the dichromats and anomalous trichromats, colour discrimination loss was correlated with colour constancy performance (i.e. the greater the discrimination loss, the lower the colour constancy).

However, when Rüttiger et al. (2001) used an achromatic adjustment task when measuring colour constancy in red-green anomalous trichromats, dichromats and normal

trichromats, they found no significant differences in colour constancy abilities between the observer groups. Furthermore, they found that for all observer groups, colour constancy performance did not seem to correlate with either luminance or colour discrimination thresholds. Despite the red-green discrimination thresholds for the red-green deficient observers being much higher than normal trichromats, their colour constancy performance was not significantly different, even when tested along three different illumination chromatic axes (red-green, blue-yellow and Planckian/daylight locus). The only difference was that for red-green deficient observers, they had higher variability along the red-green axis. Rüttiger et al. (2001) suggested that colour constancy performance may be determined, to a large extent, by the S-cone opponent mechanism (Mollon et al., 1990), which is mostly unaffected in red-green deficient observers.

The S-cone opponent mechanism, and the perception of blue to yellow hues is mostly intact in red-green deficient observers. The suggestion made by Rüttiger et al. (2001) above is potentially supported by a number of studies done on normal trichromats. These studies (e.g. Pearce et al., 2014; Weiss et al., 2017) show that for normal trichromats, there is a strong bias towards blue illuminants. In other words, when tested on illumination shifts closer to the blue skylight illumination, trichromats show much better colour constancy (decreasing the further away the illuminant shifted from CIE D65). This effect, termed the “blue bias”, suggests that colour constancy is optimized for blue skylight. Indeed, a study conducted Amano et al. (2003) finds that whilst red-green deficient observers showed elongated distributions in their colour constancy performance along the red-green axis (similar to Rüttiger et al.’s (2001) results), they also performed just as well as normal trichromats when illuminant shifts were along the daylight locus.

Therefore, one explanation for the relatively good colour constancy performance in red-green deficient observers would be that the S-cone input is preserved, and that most surfaces

illuminated under blue skylight tend to reflect more bluish colours, making S-cone input much more important for colour constancy under such circumstances. A recent study by Ma et al. (2016) supports this idea. Using a von Kries adaptation model to estimate cone responses, results indicated that colour constancy in red-green deficient observers are mediated via cone-level adaptation, in particular the S-cone opponent mechanism (especially for blue illumination that stimulate the S-cones strongly). Thus, these results by Ma et al. (2016) support the idea of the "blue bias" in colour constancy. However, it is important to note that Mat et al.'s (2016) adaptation model only observed substantial S-cone adaptation under very blue illuminations. Indeed, for blue-yellow illuminations (that were not as blue), the S-cones could only partially explain the red-green deficient observers' colour constancy performance.

Furthermore, the explanation of "blue bias" does not explain how red-green deficient observers still have relatively good colour constancy under red or green illuminants (Rüttiger et al., 2001; Amano et al., 2003). Ma et al.'s (2016) results directly contradict these previous studies because they found that dichromatic observers showed almost no colour constancy under red or green illuminants. It is unclear why these results differ, and as Ma et al. (2016) note, their study was based around assumption that dichromatic colour vision is simply a reduction of normal trichromatic colour vision. In reality, this may not exactly be the case, and more so when it comes to processing other visual cues as well as computing colour constancy.

Apart from the illuminants, surface reflectances chosen may also have a significant effect on differences in colour constancy performance between the observer groups. For example, the reflectance spectra of the surfaces chosen have shown to affect colour constancy performance for red-green deficient observers. Whilst most previous studies primarily use reflectances selected from the Munsell Book of Colour (Munsell, 1976), which contain a large gamut of uniformly distributed colours, the colours in natural scenes tend to be quite different. Indeed, Baraas et al. (2010) note that the colour gamut of natural scenes is much more

constrained compared to Munsell reflectances: natural scenes primarily contain green (grass, foliage), browns (earth, soil) and blues (water and sky).

By testing red-green deficient observers and normal trichromats on both natural and Munsell reflectances in an asymmetric colour matching task, Baraas et al. (2010) found that whilst dichromats performed significantly lower overall than normal trichromats on the tasks, their performance was significantly higher for natural reflectances than Munsell reflectances. In addition, anomalous trichromats (protanomalous and deuteranomalous) did not perform significantly different from normal trichromats, although deuteranomalous trichromats did show significantly higher colour constancy indices for natural reflectances compared to Munsell reflectances. To explain the colour constancy performance of the red-green deficient observers, Baraas et al. (2010) suggest that images of natural scenes contain less deviation noise (in terms of spatial cone-excitation ratios) which may account for better performance in the natural reflectance condition, compared to the Munsell reflectance condition (Foster et al., 2006).

The studies that have been mentioned so far comparing red-green deficient observers and normal trichromats have primarily used traditional colour constancy tasks, with an emphasis on 2D Mondrian pattern-based stimuli with Munsell reflectances: colour naming (Montag, 1994; Morland et al., 1997), achromatic adjustment (Rüttiger et al., 2001) and asymmetric colour matching (Amano, 2003; Baraas et al., 2010; Ma et al., 2016).

Whilst 2D Mondrian stimuli have been extensively used in many classic colour constancy studies due to the ability to easily isolate colour from other visual cues (Arend & Reeves, 1986; Arend et al. 1991; Maloney, 2003), it has been argued that these impoverished 2D stimuli lack the realism and visual cues normally available in the real-world, and thus colour constancy mechanisms (that operate in the real-world) may not work as effectively, and in some cases can produce conflicting results (Smithson, 2005). Therefore, some researchers have

advocated that the best way to investigate colour constancy is to use 3D objects under natural viewing conditions (Brainard et al., 2003). Studies have shown that colour constancy performance improves for 3D versus 2D scenes in normal trichromats, independent of the target colour and illuminant shift (Hendrich, Bloj & Ruppertsberg, 2009; Xiao et al. 2012).

The importance of using realistic stimuli is further emphasised when it comes to comparing colour constancy in red-green deficient observers versus normal trichromats. As red-green deficient observers have greater difficulty in discriminating between hues compared to normal trichromats, they may therefore rely more on other visual cues such as luminance or contrast (Rigden, 1999) in order to distinguish between surfaces or objects. Red-green deficient observers have shown to be better than trichromats at breaking colour camouflage and identifying the correct stimuli when presented with a series colour-camouflaged stimuli (Morgan & Mollon, 1992; Saito et al., 2006). The superior performance of red-green deficient observers on colour camouflage tasks is attributed to processing the visual scene relying on luminance texture cues. Whereas a normal trichromat processes and segregates the visual scene based on both colour and texture, dichromats are less reliant on colour cues and thus not susceptible to colour-based interference when identifying texture (Morgan & Mollon, 1992). Therefore, it is important to choose realistic and controlled stimuli which include a variety of visual cues when comparing the colour constancy performance of dichromats and trichromats.

This thesis will aim to address the underlying explanations for colour constancy performance in dichromats and trichromats: in Chapter 2, an experiment based on the literature reviewed here, comprised of comparing performance on two colour constancy tasks (2D achromatic adjustment vs. 3D blocks-copying task) will be reported. In Chapter 3, a computerised simulation of dichromatic vs. trichromatic observers will show how and what combinations of surfaces and illuminants produce the largest differences between observer types.

Table 1.1.
Experimental studies that have specifically compared the colour constancy performance of dichromats, anomalous trichromats and normal trichromats. Differences in colour constancy performance between the observer groups are notated as follows: A>B (Group A performed significantly higher than Group B), A<B (Group A performed significantly lower than Group B) and A ≈ B (No significant differences between Group A and B). Studies are tabulated in chronological order (earliest to most recent).

Study	Number and type of participants	Task type	Stimuli type	Differences in colour constancy performance
Morland et al. (1997)	9 normal trichromats 2 protanomalous trichromats 2 deuteranomalous trichromats 4 deuteranopes	Colour naming	Physical Mondrian colour boards	Anomalous trichromats < Normal trichromats Dichromats < Normal trichromats
Ruttiger et al. (2001)	5 normal trichromats 3 deuteranomalous trichromats 4 protanopes 3 deuteranopes	Simultaneous achromatic adjustment	2D Mondrian patterns	Anomalous trichromats ≈ Normal trichromats Dichromats ≈ Normal trichromats
Amano et al. (2003)	10 normal trichromats 1 protanomalous trichromat	Successive asymmetric matching	2D Mondrian patterns	Anomalous trichromats ≈ Normal trichromats

	2 protanopes 2 deuteranopes			Dichromats ≈ Normal trichromats
Baraas et al. (2006)	12 normal trichromats 7 protanomalous trichromats 7 deuteranomalous trichromats	Successive asymmetric matching	2D images of natural scenes	Deuteranomalous trichromats ≈ Normal trichromats Protanomalous trichromats < Deuteranomalous and normal trichromats
Baraas et al. (2010)	9 normal trichromats 5 protanomalous trichromats 9 deuteranomalous trichromats 5 protanopes 5 deuteranopes	Successive asymmetric matching	2D Mondrian patterns	Protanopes and deuteranopes < Normal trichromats Anomalous trichromats ≈ Normal trichromats
Ma et al. (2016)	5 normal trichromats 2 deuteranomalous trichromats 3 protanopes	Simultaneous haloscopic colour matching	2D Mondrian patterns	Dichromats < Normal and deuteranomalous trichromats

	1 deuteranope			
Alvaro et al. (2017)	4 normal trichromats 3 protanopes 4 deuteranopes	Successive asymmetric forced-choice matching	2D images of natural scenes (urban and rural)	Dichromats ≈ Normal trichromats
Pastilha et al. (2019)	4 normal trichromats 2 protanopes 2 deuteranopes	Successive asymmetric colour-object matching task (forced choice)	Real 3D geometric objects	Dichromats < Normal trichromats (but higher than estimated)

Chapter 2: Proposed Experiment

Due to COVID-19, access to the vision laboratory and participant recruitment was restricted and proceeding with the experiment was not possible. The experiment did not take place as originally planned, due to the nature of this experiment involving colour-vision-deficient observers, as well as it being not possible to conducting colour constancy tasks on online platforms (e.g. colour calibration of participants' monitors, accurate screening/measurements for colour vision deficiencies). Below are the detailed aims, design, and methods and procedure/tasks that would have been used for the proposed experiment. Prior to COVID-19, the experimental design was planned, the subject groups identified, and the proposed study was approved by the Durham University Department of Psychology Ethics Committee.

For ease of reading, this chapter is written in the past tense, however, the experiment has not actually been conducted and no data has been collected: therefore, this chapter will not have a results or discussion section.

2.1 Aims and hypotheses

This experiment aimed to investigate how colour constancy performance in trichromatic and red-green dichromatic observers differ under task and illuminant type. Previous studies have been rather inconclusive as to the extent and nature of colour constancy in colour vision deficient observers (e.g. Rüttiger et al., 2001; Baraas et al., 2010; Ma et al., 2016; Alvaro et al., 2017). Limitations with previous studies are to be addressed in this study – namely, these previous studies have only compared colour constancy performance using either 2D or 3D stimuli, and always with only one task type – see Table 1 in Chapter 1. As

mentioned in the review, colour constancy performance improves when stimuli are 3D compared to 2D (Hendrich, Bloj & Ruppertsberg, 2009; Xiao et al. 2012).

In addition, another problem regarding traditional 2D tasks is that colour vision deficient observers may rely more on non-colour visual cues such as luminance or contrast (Rigden, 1999) in order to distinguish between surfaces or objects. Therefore, colour constancy tasks relying purely on 2D Mondrian stimuli as well as adjustment-type tasks (rather than selection-type tasks) may not best represent their colour constancy performance in real-life.

Thus, the aim of this study is to address the limitations of previous studies, as well as provide a new avenue for colour constancy/colour vision deficiency research by explicitly comparing the colour constancy performance of trichromats and dichromats on two task types: a traditional 2D achromatic adjustment task, and a 3D blocks-copying task. In addition, the illumination and the so called "blue-bias" and its effect on colour constancy will also be investigated with regards to red-green anomalous trichromats and dichromats, as little is known about it (Pearce et al., 2014; Weiss et al., 2017).

Based on the research done in this area so far, it is therefore predicted that:

1. Within each group, both trichromats and dichromats will perform significantly better on the 3D blocks-copying task than the 2D achromatic adjustment (between groups, trichromats will perform better than dichromats on average).
2. Dichromats and trichromats will perform significantly worse on both colour constancy tasks when the illuminants are orthogonal to daylight locus, due to the illuminants not being part of the daylight illuminants that colour constancy is tuned for ("blue-bias"). In addition, dichromats and anomalous trichromats will perform significantly more variable when illuminants are orthogonal to the daylight locus (due to these illuminants falling upon their confusion axes), but their mean

constancy performance will remain similar to normal trichromats (e.g. Amano et al., 2003).

3. There will be an interaction effect between the level of CVD and task type. The difference in 3D vs 2D colour constancy task performance will be greater for dichromats than for trichromats (because dichromats are more dependent on 3D non-colour visual cues than trichromats).

2.2 Design

The experiment involved measuring the performance of dichromatic, trichromatic and anomalous trichromatic observers using two different colour constancy tasks in a 2 x 2 factorial design. The experiment used a between-subjects design. The following independent variables were manipulated, each with two levels:

1. The colour constancy task type (traditional 2D achromatic adjustment task vs. naturalistic 3D blocks-copying task).
2. Illuminant type (daylight locus vs. orthogonal to daylight locus).

These tasks will be described in more detail in Section 2.4. The dependent variable was the colour constancy performance of the participant, measured and calculated using a colour constancy index.

2.3 Apparatus and stimuli

The stimuli were displayed on a colour-calibrated and gamma-corrected monitor. A Cambridge Research Systems ViSaGe MKII colour stimulus generator was used to generate and display the stimuli. Responses were taken using a regular Microsoft keyboard and mouse for both tasks.

The tasks and visual stimuli for this experiment were was programmed in MATLAB with the help of the MATLAB toolbox, Psychtoolbox-3. Data was collected within the same script.

2.4 Procedure/Task

Participants were recruited via various online platforms as well as flyers posted around on the Durham University campus. Prior to COVID-19, the proposed study was approved by the Durham University Department of Psychology Ethics Committee. All participants were required to give informed consent and were free to withdraw anytime during the study. In addition, all participants were told that their data would be anonymised afterwards upon agreeing to take part.

After participant recruitment, screening took place. Firstly, all participants were made sure to have normal or corrected-to-normal visual acuity. Secondly, colour blindness tests were conducted in order to determine whether the participant had a colour vision deficiency. For initial screening of red-green dichromats, the Ishihara pseudo-isochromatic test was used. To categorise the other observer types (normal trichromats and anomalous trichromats), Rayleigh anomaloscope measurements were taken and analysed.

After the screening, participants were divided in one of three groups based on their colour vision deficiency – normal trichromat, red-green anomalous trichromat or red-green dichromat.

Participants then proceeded to complete two colour constancy tasks. To avoid fatigue effects, participants completed both tasks on two separate days, and the order in which the task was done was randomised. Each task took no more than 30 minutes to complete, in two 15-minute blocks with a break in between. The two colour constancy tasks were:

1. A 2D traditional achromatic adjustment task

2. A 3D naturalistic blocks-copying task

An example of the stimuli presented to the observers are presented are below (Figure 2.1 and 2.2).

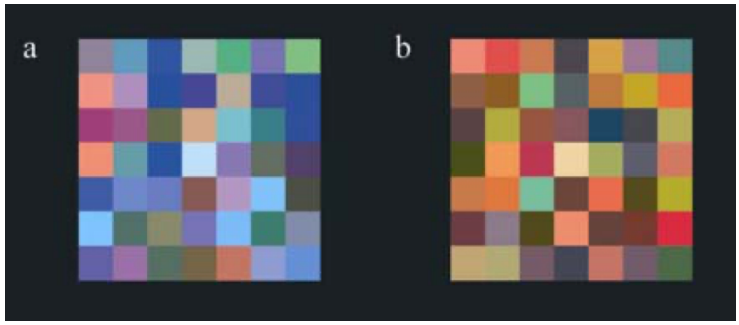

Figure 2.1. An achromatic adjustment task involves asking the observer to adjust the central patch until it appears white. Pattern A is illuminated by blue skylight while pattern B is illuminated by red sunlight. The two central patches show a successful adjustment to white. Source: Foster (2003).

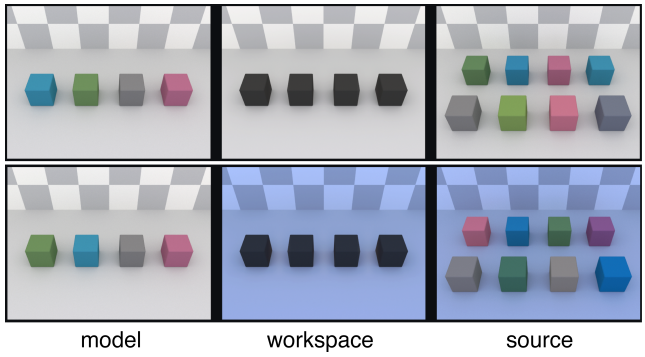

Figure 2.2. Three rendered rooms with simple block-like objects are presented to the participant. Subjects are asked to recreate the arrangement of blocks in the model by replacing the 4 blocks in the workspace from a selection of 8 blocks from the source. The top row shows the same illuminant for all three rooms, whilst the bottom row shows an illuminant change for the workspace and source. Source: Radonjić, Cottaris & Brainard (2015).

Alongside the task type, illuminants were manipulated, with a set of illuminants along the daylight locus versus a set of illuminants orthogonal to the daylight locus. In addition, a selection of 30 different Munsell reflectances were presented. 10 unique Munsell hues with medium lightness, alongside three levels of chroma (saturation) were specifically chosen as the

surface reflectances to be tested (for more information, see Chapter 3). The surface reflectance and illuminant selections were kept constant across both tasks.

In the 2D achromatic adjustment task (Figure 2.1), participants were presented a 2D Mondrian pattern with a central patch under a certain illuminant, and their task was to adjust the colour of the central patch using four buttons on a keyboard, which corresponded to the four cardinal colour directions in CIE L* a* b* colour space (red, green, blue and yellow). The participants' final aim was to ensure that the central patch looked "white" to them under that specific illuminant.

In the 3D naturalistic blocks-copying task (Figure 2.2), participants were presented three 3D rendered rooms (model, workspace and source) and were asked to use a mouse to select and drag the and match four blocks (from a selection of 8) in the source into to the workspace. The aim was to select the four blocks in the source which most matched the surface reflectance of blocks in the model. The model room would be presented under the neutral illuminant, whereas the workspace and source would be presented under a different illuminant.

For both tasks, there was no time-limit per trial, and participants were encouraged to take as long as they needed to adjust/match the surfaces. Once their choice was finalised, the next trial would be presented. In addition, to ensure the participant understood the aim of the task and were not simply matching the hue-saturation of the coloured surfaces, several demo trials were presented before the actual task began, and participants were asked to repeat back the goal of the task in their own words to the researcher.

Chapter 3: Computerised Simulation

3.1 Aims and Design

This part of the project involves a computerised simulation of trichromatic vs. dichromatic and anomalous observers. This simulation is aimed to model ideal dichromatic, trichromatic and anomalous trichromatic observers when presented with various surfaces and illuminants.

The following research questions form the basis of this simulation:

1. Do trichromats, dichromats and anomalous trichromats differ in terms of colour constancy, and if so, how?
2. How is the discriminability of surfaces under different illuminants affected by observer type? In other words, what combinations of surfaces and illuminants produce the largest differences between observer types?

The specific surface-illuminant combinations that produce the largest differences between observer types are of particular interest here. The results from this computerised simulation are used to decide which surfaces and illuminants to include in the colour constancy tasks in the experiment in Chapter 2. Furthermore, differences in the colour constancy performance of the ideal observer models in this computerised simulation and colour constancy performance in the experiment can be compared – this will provide insight into how colour constancy mechanisms in trichromats, anomalous trichromats and dichromats work in real life.

3.2 Methods

By using a set of fixed Munsell reflectances, a variety of five different illuminants, and using the DeMarco cone fundamentals (DeMarco et al., 1992), the cone excitation values for each separate cone type can be calculated. This is basically akin to the reflected light/colour signal at the cones (Reflectance × Illuminant × Cone sensitivity). All spectral data in this simulation were eventually sampled (splined) from 400 to 700nm, with a wavelength resolution of 5nm. There are four variables of interest in this simulation: surface reflectances, illuminants, observer type and colour constancy measure. This will be explained in detail below:

1. **Surface reflectances**. 1269 surface reflectances for matte patches in the Munsell Book of Color were downloaded from https://www.uef.fi/web/spectral/munsell-colors-matt-spectrofotometer-measured. The reflectance spectra were measured with a Perkin-Elmer lambda 9 UV/VIS/NIR spectrophotometer, with a wavelength resolution of 1nm. The standard Munsell notation was used in labelling and classifying the surfaces: hue, value and chroma (e.g. 5R 5/6). For this simulation, a subset of 30 (out of the 1269) surfaces were chosen. These 30 surfaces were chosen on the basis of their chromas. The hues and values were kept constant, or as similar as possible. The 30 surfaces were then divided into three surface sets: low, medium and high chroma. See Table 3.1 below.

Table 3.1.
Subset of Munsell reflectances chosen for the computerised simulation. Each surface set contains 10 different Munsell reflectances.

Surface set	Surface set property	List of Munsell reflectances used
1	Low chroma	5R 5/1, 5YR 5/1, 5Y 5/1, 5GY 5/1, 5G 5/1, 5BG 5/1, 5B 5/1, 5PB 5/1, 5P 5/1, 5RP 5/1
2	Medium chroma	5R 5/6, 5YR 5/6, 5Y 5/6, 5GY 5/6, 5G 5/6, 5BG 5/6, 5B 5/6, 5PB 5/6, 5P 5/6, 5RP 5/6
3	High chroma	5R 5/8, 5YR 5/8, 5Y 6/8, 5GY 5/8, 5G 5/8, 5BG 5/8, 5B 5/8, 5PB 5/8, 5P 5/8, 5RP 5/8

2. **Illuminants**. Five different illuminants were used. These were: neutral D65 daylight, yellow, blue, green and red illuminants. The neutral D65 illuminant was used as the baseline illuminant, with two daylight illuminants (yellow and blue) and two atypical chromatic illuminants (green and red) are used as the independent variables for colour constancy calculations. The illuminant spectra for the five illuminants are shown in Figure 3.1 below.

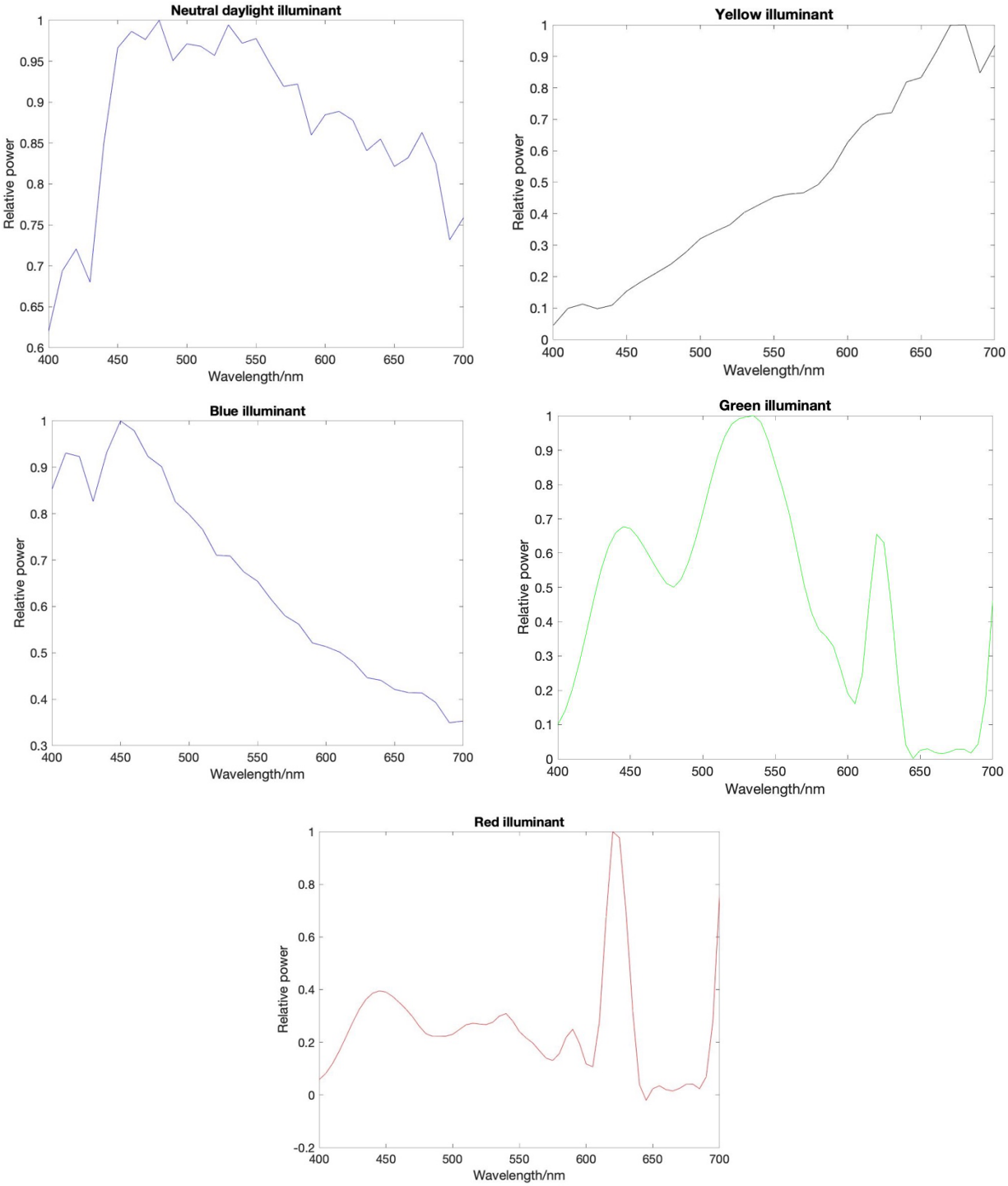

Figure 3.1. The spectra of the five illuminants that were used in this simulation. The neutral daylight illumination was used as the baseline illuminant and the four illuminants (two daylight, two atypical) were used as the independent variables. The correlated colour temperature (CCT) of the neutral daylight, yellow and blue illuminants are 5700K, 2800K and 10000K respectively.

As it can be seen from the illuminant spectra plots, the green and red illuminants are meant to represent atypical and extreme conditions – it would be rare to find natural lighting conditions with such spectra, compared to the neutral, yellow and blue illuminants. However, for the purposes of this simulation, these illuminants will be very useful in testing and comparing the limits of colour constancy performance between trichromats, anomalous trichromats and dichromats.

3. **Observer type**. There were five different observer types in this simulation: 1. Normal trichromat (LMS), Protanope (MS), Deuteranope (LS), Protanomalous trichromat (L'MS), Deuteranomalous trichromat (LM'S). To generate the ideal observer model, optimal colour spaces (i.e. perceptual dimensions) were calculated, in which quantifying changes in discriminability would be most optimal for each observer type. This is heavily based on work done by Aston et al. (2016). The rationale behind this ideal observer model is based on research done by Boehm et al. (2014) suggesting that changes in post-receptoral gain optimize post-receptoral processing of colour signals in anomalous trichromats. Anomalous trichromats maintain a trichromatic colour space but their dimensions differ to that of normal trichromats (Bosten et al., 2005).

To start at the beginning, Ruderman et al. (1998) performed principal components analysis (PCA) on a set of L, M, and S values obtained from hyperspectral images of natural scenes and found that the principal axes of variation were a luminance axis (L+M+S) and two chromatic axes (L-M and S-(L+M)), suggesting that colour processing channels have evolved to optimally trade-off the cone responses for maximum discrimination. These axes (luminance and chromatic) and PCA will form the basis of the ideal observer model detailed below.

Thus, this model is based on the assumption that post-receptoral colour processing channels may adapt in colour vision deficient observers to optimally represent the available information from the remaining and/or anomalous cone types.

Using the DeMarco cone fundamentals (DeMarco et al., 1992), the cone excitation values for each separate cone type (L, M, S, L', and M') can be calculated. This is basically akin to the reflected light/colour signal at the cones, generated by point-by-point multiplication of the reflectance spectra, illuminant spectra and cone sensitivity functions (Reflectance × Illuminant × Cone sensitivity).

The cone excitation values calculated were then scaled using von Kries adaptation (i.e. assuming the observer has perfect knowledge of the cone values of the illumination) within each cone class (L/L', M/M', S) to add an element of colour constancy to the model (von Kries, 1878).

To obtain the optimal colour space for each observer type, principal components analysis (PCA) was done on the corresponding von Kries scaled cone excitation values for all the Munsell surface reflectances. In other words, the components produced by the PCA are the axes of the optimal colour spaces. The recovered coefficients used to transform stimulus values from cone space to optimal colour space for each observer type are given in Table 3.2, where each axis has a category. For normal trichromats and anomalous trichromats, it can be seen that optimal colour space is three dimensional and one of the axes is always a luminance axis (Lum) and the other two are chromatic axes (S-(L+M) and L-M). Whereas for dichromats, their optimal colour space is two dimensional with a luminance axis and one chromatic axis.

In order to express a chosen surface in this colour space, the appropriate cone responses are computed (depending on the observer type). Then, those cone responses are von Kries scaled by an estimate of the chosen illuminant, and finally the cone responses are transformed to the optimal colour space by multiplying them with the appropriate set of coefficients. See Table 3.2 below.

Table 3.2.
Coefficients for transforming from von Kries scaled cone responses to optimal colour space for each observer type.

	Coefficients			
Observer Type	**Cone 1**	**Cone 2**	**Cone 3**	**Axis Name**
Normal Trichromat	0.61	0.61	0.51	Lum
(Cones 1-3 are L, M, S)	-0.43	-0.28	0.86	S-(L+M)
	-0.67	0.74	-0.09	L-M
Protanope	0.73	0.68	N/A	Lum
(Cones 1-2 are M, S)	-0.68	0.73	N/A	S-M
Deuteranope	0.74	0.67	N/A	Lum
(Cones 1-2 are L, S)	-0.67	0.74	N/A	S-L
Protanomalous	0.61	0.61	0.51	Lum
(Cones 1-3 are L', M, S)	-0.42	-0.30	0.86	S-(L+M)
	-0.67	0.74	-0.08	L-M
Deuteranomalous	0.61	0.61	0.50	Lum
(Cones 1-3 are L, M', S)	-0.40	-0.31	0.86	S-(L+M)
	-0.68	0.73	-0.06	L-M

4. **Colour constancy measure**. The first three variables covered so far (surface reflectances, illuminants and observer types) are the independent variables of this simulation. The dependent variable is the colour constancy measure. In order to calculate the colour constancy measure, each surface is first represented in 2D space for each observer and illuminant combination, with the axes being the most significant dimensions of the PCA solution. As an example, Figure 3.2 shows the plots for the trichromat observer with Surface Set 2 (medium chroma).

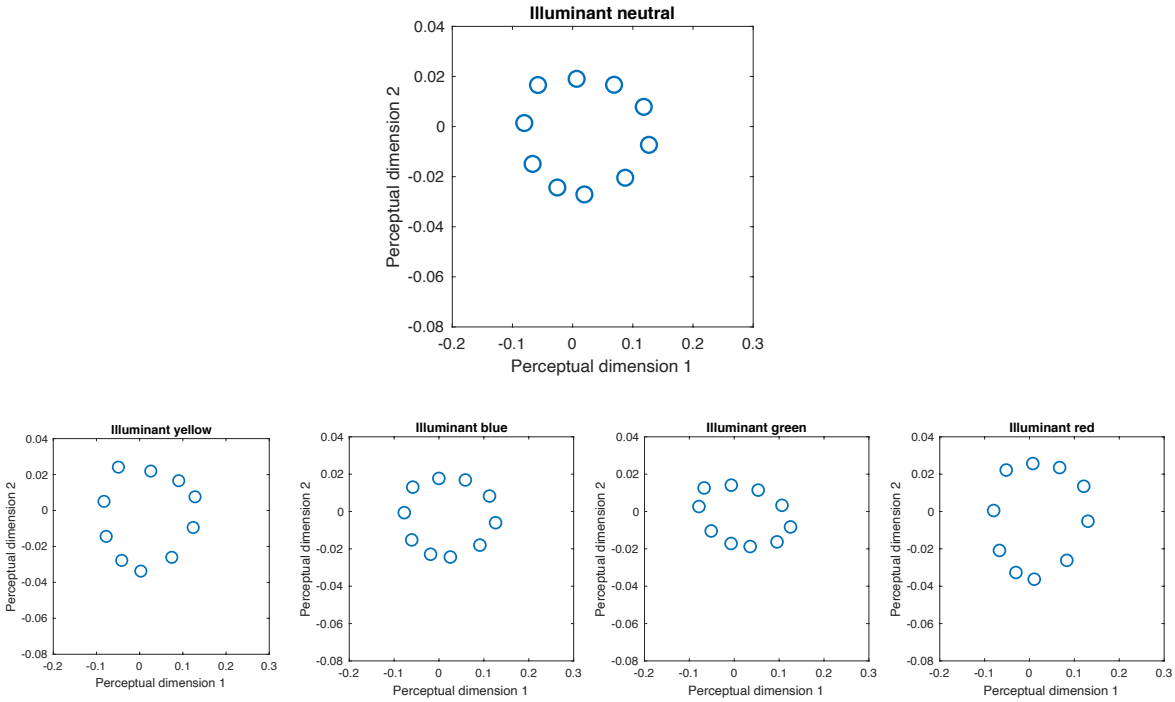

Figure 3.2. Surfaces in surface set 2 (medium chroma) for the normal trichromat observer. Each graph shows a different illuminant, with the neutral daylight illuminant (top) being the baseline measure. Each surface is represented by an open circle, and each graph contains the same 10 surfaces. Perceptual dimension 1 refers to the S-(L+M) axis, and perceptual dimension 2 refers to the L-M axis as mentioned in the previous section.

As it can be seen from the plots above for the normal trichromat observer, there are changes in the surface centroids under different illuminants, and the same can be shown for five different observer types (not included here). These same graphs can be plotted for all five observer types (normal trichromat, protanope, deuteranope, protanomalous trichromat and deuteranomalous trichromat) as well as all three surface sets. The centroids can be helpful in visualising changes in identification and discriminability within the observer.

To calculate the colour constancy measure, the Euclidean distance is calculated from each pair of points compared to the neutral illuminant. In Figure 3.2 above, each open circle that represents the surface under the yellow, blue, green and red illuminant will be compared to the same surface under the neutral illuminant. The Euclidean distances can then be calculated for all 10 surfaces and then averaged via the arithmetic mean – and this Mean Absolute Distance is the error measure with which colour constancy will be measured with.

The larger the error/distance, the greater the change in identification, meaning lower colour constancy. Therefore, when interpreting results, a lower Euclidean Distance means higher colour constancy. For graphs of the inverse colour constancy measure (Euclidean Distance), see Figure 3.3 in the results section below.

3.3 Results

A three-way ANOVA was conducted across all surfaces to examine the effect of surface set, illumination and observer type on colour constancy (as inversely measured by Euclidean Distance). The groups were found to have significantly different variances via Levene's test, $F(59, 540) = 10.915$, $p < .001$. Therefore, the results from this three-way ANOVA should be taken with caution.

Initial analysis showed that all three independent variables had significant main effects: illumination, $F(3, 540) = 68.663$, $p < .001$, surface set, $F(2, 540) = 247.059$, $p < .001$, and observer type, $F(4, 540) = 7.445$, $p < .001$. Further analysis showed that there was no statistically significant three-way interaction, $F(24, 540) = 0.132$, $p = 1.000$. However, there was a significant two-way interaction between illumination and surface set on colour constancy, $F(6, 540) = 12.240$, $p < .001$. The other two-way interactions were not significant.

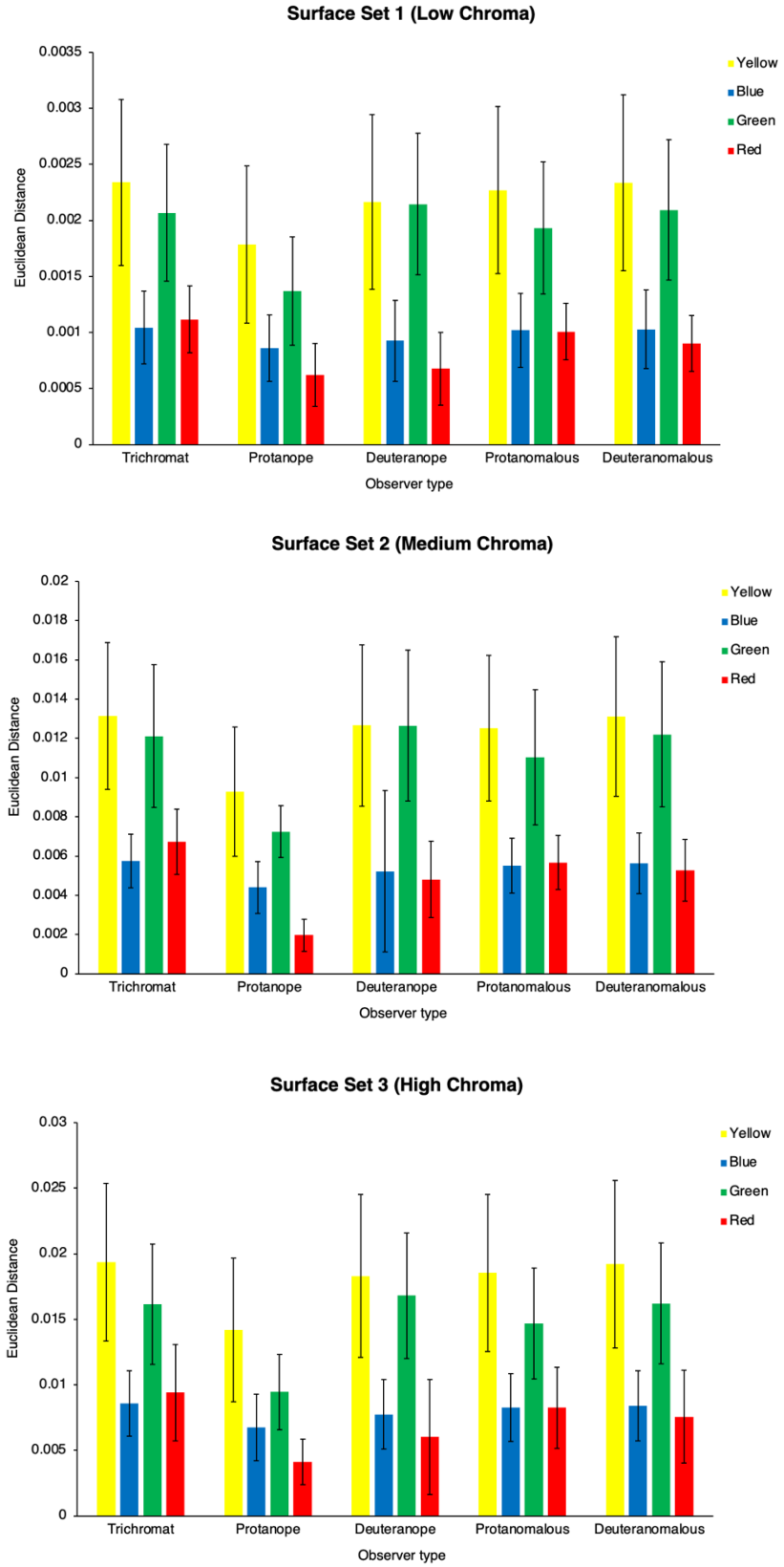

Figure 3.3. Euclidean Distance as inverse colour constancy measure (*y*-axis) plotted for the five different observer types (*x*-axis). The different coloured bars represent the illuminant shift (from neutral to either yellow, blue, green or red chromatic illuminant). Separate plots are shown for each surface set (Low, Medium and High chroma). Error bars represent 95% confidence intervals (calculated over $n = 10$). Note: the values and scaling of the y-axis are different for all three plots (to preserve visual clarity).

1. **Illumination.** Tukey post-hoc tests showed that across all observer types and surface

sets, colour constancy was significantly lower under yellow illumination compared to blue ($p < .001$), green ($p < .05$) and red ($p < .001$). Colour constancy was significantly higher under blue illumination compared to green ($p < .001$) and was also significantly higher under red illumination compared to green ($p < .001$).

2. **Surface set.** Tukey post-hoc tests showed that across all observer types and illuminants, colour constancy was significantly higher for Surface set 1 (Low chroma surfaces) compared to Surface set 2 (Medium chroma) ($p < .001$). Colour constancy was significantly higher for Surface set 1 compared to Surface set 3 (High chroma) ($p < .001$). Finally, colour

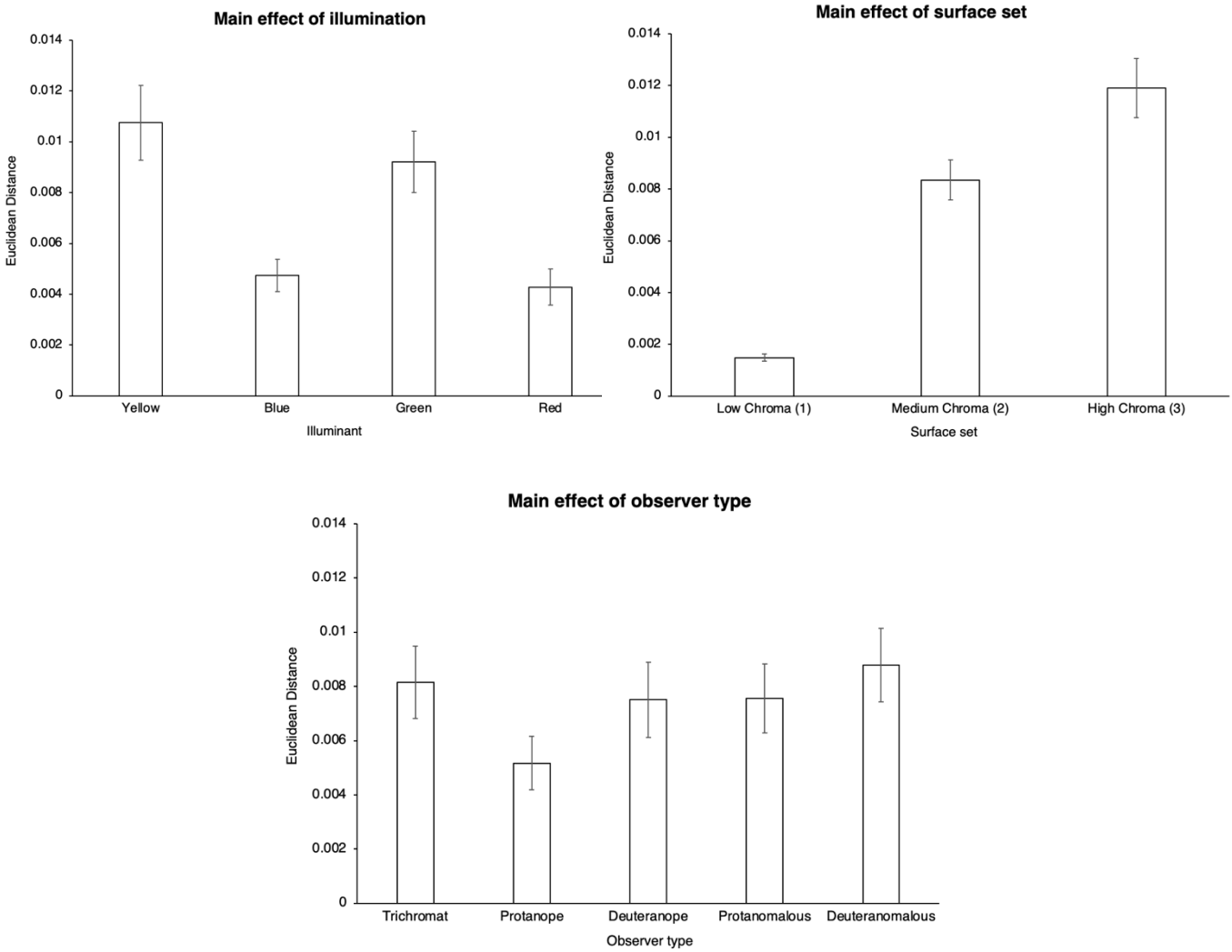

Figure 3.4. Main effect of the three independent variables on inverse colour constancy measure (Euclidean distance). *y*-axes on all three graphs are the same for easier visual comparison. Error bars shown are 95% confidence intervals (calculated over n = 150, 200 and 120 for illuminant, surface set and observer type respectively).

constancy was significantly higher for Surface set 2 compared to Surface set 3 ($p < .001$). The two-way interaction between illumination and surface set was further analysed and colour constancy for surface set 1 (across all illuminants) was found to be significantly higher than surface set 2 and 3 ($p < .001$). Colour constancy was also higher for surface set 2 compared to 3, but only for yellow and blue illuminants ($p < .05$).

3. **Observer type.** Tukey post-hoc tests showed that across all surface sets and illuminants, colour constancy was significantly higher for protanopes compared to all other observer types ($p < .01$). Colour constancy for trichromats, deuteranopes, protanomalous trichromats and deuteranomalous trichromats were not significantly different.

3.4 Discussion

The section will interpret the results of the three-way ANOVA, whilst simultaneously discussing the methodological limitations of the current simulation. It is important to reiterate that the results of the three-way ANOVA have to be taken with caution due to the groups not having homogeneity of variances. Results showed that there was no significant three-way interaction, but there were significant main effects for all three independent variables.

1. **Illumination.** The first significant main effect was illumination. Across all observer types and surface sets, the main effect of illumination can be summarised as follows:

1. Colour constancy was significantly lower under the yellow illuminant compared to the rest of the illuminants (blue, green, red).
2. Colour constancy was significantly higher under blue illumination compared to green.
3. Colour constancy was also significantly higher under red illumination compared to green.

In this simulation, colour constancy was measured via an illuminant shift from a neutral D65 daylight to one of four target illuminants (yellow, blue, green or red). The target

illuminants were either daylight (yellow and blue) or atypical illuminants (green and red). In terms of chromatic illuminant shifts, Worthey (1985) showed that colour constancy is better for "blue–yellow" than "red-green" illuminant shifts, which is supported by Mollon et al. (1990) and Rüttiger et al. (2001), suggesting that colour constancy performance may be heavily dependent on the S-cone opponent mechanism. A more recent study by Daugirdiene et al. (2016) found that constancy was higher for shifts from a neutral illuminant to an extreme blue illuminant, compared to yellow. On the other hand, Wan and Shinomori (2018) found the opposite: that constancy was higher for red-green illuminant shifts compared to blue-yellow, with lowest constancy for blue illumination shifts.

In the current simulation, results partially agree with Wan and Shinomori (2018): with the constancy being lowest under yellow illuminant shift, but also disagree, with constancy for the blue illuminant shift higher than green. Higher constancy for blue illuminants is supported by literature on the blue-bias of colour constancy (Pearce et al., 2014; Weiss et al., 2017). However, it is unsure why constancy was higher under red compared to green.

Nevertheless, the main effect of illumination on colour constancy in this simulation should be interpreted with caution. The simulation did not involve actual observers doing a task, and the observers' were modelled based on von Kries adaptation (involving perfect knowledge of the cone scaling values of the illuminant). Indeed, Hulbert (2019) states that the difference in results between Daugirdiene et al. (2016) and Wan and Shinomori (2018) can be explained by the difference in the adaptation state of the observer (involving haplosopic viewing, where both eyes are exposed separately and simultaneously to different illuminants). Therefore, it is clear that in this simulation, where von Kries adaptation modelling is used, results may differ significantly compared to an actual colour constancy task.

Finally, the colour constancy measure here was compared across all observer types (including the red-green deficient observers), with no interaction effect between illumination

and observer type. Thus, this means that under these four illumination shifts (and three surface sets), the colour constancy of red-green anomalous trichromats or dichromats was found to not be significantly different from normal trichromats. This is similar to the result found by Rüttiger et al. (2001), where they also tested constancy changes along the three illumination shifts of red-green, blue-yellow and daylight locus. As the S-cone opponent mechanism, proposed to be crucial in colour constancy (Mollon et al., 1990) is mostly unaffected in red-green deficient observers, the non-significant interaction effect between illumination and observer type on colour constancy found in this simulation directly supports the conclusions made by Mollon et al. (1990) and Rüttiger et al. (2001). The simulation here also contradicts the adaptation modelling done by Ma et al. (2016), who found that dichromats showed almost no constancy under red or green illuminants.

2. **Surface set.** The second significant main effect was surface set. Across all observer types and illuminants, results indicated that colour constancy was significantly higher, across all illuminants and observer types, for surface set 1 (Low chroma surfaces) compared to surface set 2 (Medium chroma) or 3 (High chroma). Colour constancy for surface set 2 was also significantly higher than surface set 3.

The Munsell colour system, which each of the surfaces in this simulation are defined by, are defined by three separate and independent values (Tyler & Hardy, 1940; Nickerson, 1940; CIE, 2020). Hue (colour attribute: red, yellow, green, blue and purple), value (lightness, from pure white to pure black), and chroma (saturation of the surface judged/compared to a similarly illuminated area that appears white). With the hues and values being kept constant, one of the variables manipulated here was chroma.

Overall, this simulation found that there was a strong negative relationship between chroma and colour constancy, with observers exhibiting highest colour constancy when the surfaces had a lower level of chroma. One explanation for this would be that surfaces with a

higher level of chroma (increase in vividness) means that the illuminant shift has less of an effect on the overall colour appearance for the observer. In other words, when shown surfaces with higher levels of chroma, observers are less able to discriminate whether the material colour actually changed under the neutral D65 vs. chromatic illuminant. It is also important to point out that this simulation used measurements of actual matte Munsell surfaces, and all the colours can exist on real life surfaces. However, an alternative explanation for this result could be that low chroma surfaces have low variance since all surfaces have similar reflectance functions, whereas high chroma surfaces have more variation in their spectral distribution and would therefore naturally have larger changes with under the different illuminants.

The finding that highly chromatic surfaces can lower colour constancy performance is supported by Nascimento et al. (2004) and Morovič and Morovič (2005). From both a theoretical and practical perspective, colour constancy is inextricably tied to the chroma of the surface, and highly chromatic surfaces cannot be colour constant (Morovič & Morovič, 2005). Thus, the current findings of effects of surface chroma on constancy are unsurprising.

3. **Observer type.** Across all illuminants and surfaces, colour constancy was significantly higher for protanopes compared to all other observer types. However, colour constancy for trichromats, deuteranopes, protanomalous trichromats and deuteranomalous trichromats were not significantly different. This finding is interesting because all of the studies investigating red-green dichromats have either come to the conclusion that colour constancy should be similar or worse in red-green dichromats compared to normal trichromats (Ruttiger et al., 2001; Amano et al., 2003; Baraas et al., 2006, Ma et al., 2016; Alvaro et al., 2017; Pastilha et al., 2019). This simulation, however, found the reverse.

One explanation for this result would be the observer model used in this simulation, coupled with the colour constancy measure used. Firstly, the observer model used was based on optimal colour spaces, with dichromats lacking a third dimension of colour vision compared

to trichromats. This naturally means that dichromats have their colour signals restricted to one axis, making them less sensitive to changes in certain colours. When expressing the surface reflectance and illuminant as a cone signal, information from both variables is lost and compressed (compared to normal trichromats).

When calculating the colour constancy measure, a Euclidean distance/error measure was used. The higher the distance/error, the lower the constancy. It is perhaps this combination of the natural decrease in variance of chromaticities (due to how the observer model was derived) and the choice of surfaces/illuminants of the protanope that led to significantly higher colour constancy. However, deuteranopes (who lack the M cone) have a similar decrease in variance of chromaticities compared to protanopes (who lack the L cone). If the higher colour constancy performance of protanopes is due to the natural decrease in Euclidean error, this should be somewhat reflected in deuteranopes as well. Despite this, results showed that deuteranopes' colour constancy was significantly lower than protanopes.

Overall, this simulation found some differences in colour constancy across illuminants, surface set (chroma) and observer type. It is important to be cautious about interpreting the results from this simulation, as several factors have not been properly addressed here. The first factor is the observer models. The ideal observer models used in this simulation were entirely cone-based, and von Kries adaptation (von Kries, 1878) was applied to the cone models. In reality, von Kries adaptation is a good approximation of what happens at the cone level, but observers do not have perfect knowledge of the illuminant spectra – and there is ample evidence to show that von Kries adaptation is limited in its scope to fully explain colour constancy (Worthey & Brill, 1986; Kulikowski et al., 2012). In addition, cone-based models like the one used in this simulation are an oversimplification of the colour constancy process in real life – ignoring the cortical mechanisms (Smithson, 2005; Barbur & Sprang, 2008; Bannert & Bartels, 2017).

Future research could focus on comparing results from cone-based simulations, like the one here, to experimental data from actual observers doing colour constancy tasks. For example, by comparing the colour constancy measure of observer models in this simulation to experimental data from the proposed experiment in Chapter 2, differences in colour constancy performance can be interpreted and attributed to non-cone level colour constancy mechanisms, such as the different cognitive mechanisms required in completing the traditional vs. naturalistic colour constancy tasks.

The second factor is the way the simulation was run and how the data was analysed. The simulation was essentially run over a single observer, repeating it over all the 600 surface/illuminant/observer combinations. This meant that the colour constancy measure reflected variability over the different surfaces, meaning variability might have been higher than expected. An improvement to this approach would be to run this simulation *n* times, where *n* is the number of modelled observers, with each modelled observer having individual parameters and noise. This would more accurately reflect the variability across subjects, similar to how an actual experiment is conducted.

Chapter 4: General discussion

This thesis has extensively focused on the specific problem of colour constancy in dichromats and trichromats. After reviewing the current studies in this area, an experiment comprised of comparing performance on two colour constancy tasks (2D achromatic adjustment vs. 3D blocks-copying task) was proposed and reported here, and a computerised simulation showed interesting differences in colour constancy across a set of 30 matte Munsell surfaces with differing chroma, four chromatic atypical illuminants, and the five different observer types. The following sections will provide a short summary of the overall limitations of the current research done here, and what future research can be done.

4.1 Limitations of current research

A major portion of the research was significantly hindered by the COVID-19 pandemic. Thus, the scope of this thesis has been altered accordingly, with an extended literature review section (Chapter 1) and the simulation (Chapter 3) forming a larger part of the thesis. Despite this, the initial research plan was primarily experimental and involved primary data collection and analysis from conducting the proposed experiment in Chapter 2. Carrying out the experiment and recruiting participants was not possible. Thus, the results from the computerised simulation (Chapter 3) could not be compared to experimental data (Chapter 2), which would have yielded further insight into the nature of colour constancy in dichromats vs. trichromats. Specifically, how task type (2D vs. 3D, adjustment vs. selection-based task) affect colour constancy could not be addressed.

In addition, the simulation was designed to support the experiment, and in itself cannot fully address the research questions. However, the simulation did find significant main effects on colour constancy performance across surfaces, illuminants, and observer type. These results,

however, have to be taken with caution due to how the simulation was run once for each observer type repeated across all surfaces (ignoring the variability between individual observers of the same type), as well as the assumptions behind the observer model (cone-based von Kries adaptation model).

4.2 Summary and future research

So far, whilst there has been much research into colour constancy performance of normal trichromats, there are only a handful of studies that have compared colour constancy of normal trichromats with dichromats and anomalous trichromats (Ruttiger et al., 2001; Amano et al., 2003; Baraas et al., 2006, Ma et al., 2016; Alvaro et al., 2017; Pastilha et al., 2019), and results have been inconclusive, with some showing that the constancy of red-green deficient observers are just as good as normal trichromats, and other studies showing that constancy of red-green deficient observers are worse than normal trichromats.

The proposed experiment aimed to address the limitations of previous studies (namely, the use of only one task and a small number of participants) by explicitly comparing the colour constancy performance of trichromats and dichromats on two task types: a traditional 2D achromatic adjustment task, and a 3D blocks-copying task. For future research, this experiment could be performed on a large number of participants. This would provide insight into the mechanisms underpinning colour constancy in real life, as well as how constancy for dichromats and trichromats is dependent on task type.

Finally, the computerised simulation could be further expanded upon by modelling individual subjects, with each subject having individual parameters and noise. This more complex approach would more accurately reflect the variability across subjects, allowing for a better model of colour constancy at the cone level.

References

Álvaro, L., Linhares, J. M., Moreira, H., Lillo, J., & Nascimento, S. M. (2017). Robust colour constancy in red-green dichromats. *PloS one*, *12*(6), e0180310.

Álvaro, L., Moreira, H., Lillo, J., & Franklin, A. (2015). Color preference in red–green dichromats. *Proceedings of the national academy of sciences*, *112*(30), 9316-9321.

Amano, K., Foster, D. H. & Nascimento, S. M. C. (2003). Red-green colour deficiency and colour constancy under orthogonal-daylight changes. *Normal and Defective Colour Vision*, 225-230.

Arend, L. E., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: papers with diverse Munsell values. *JOSA A*, *8*(4), 661-672.

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *JOSA A*, *3*(10), 1743-1751.

Aston, S., Turner, J., Bisson, T. L. C., Jordan, G., & Hurlbert, A. (2016). Better colour constancy or worse discrimination? Illumination discrimination in colour anomalous observers. In *39th European Conference on Visual Perception (ECVP)*. Newcastle University.

Bannert, M. M., & Bartels, A. (2017). Invariance of surface color representations across illuminant changes in the human cortex. *Neuroimage*, *158*, 356-370.

Baraas, R. C., Foster, D. H., Amano, K., & Nascimento, S. M. (2010). Color constancy of red-green dichromats and anomalous trichromats. *Investigative ophthalmology & visual science*, *51*(4), 2286-2293.

Barbur, J. L., & Spang, K. (2008). Colour constancy and conscious perception of changes of illuminant. *Neuropsychologia*, *46*(3), 853-863.

Barbur, J. L., & Rodriguez-Carmona, M. (2017). Colour vision requirements in visually demanding occupations. *British medical bulletin*, *122*(1), 51-77.

Barbur, J., Rodriguez-Carmona, M., Evans, S., Authority, C. A., & Milburn, N. (2009). Minimum Color Vision Requirements for Professional Flight Crew, Part III: Recommendations for New Color Vision Standards.

Berson, D. M. (2007). Phototransduction in ganglion-cell photoreceptors. *Pflügers Archiv-European Journal of Physiology*, *454*(5), 849-855.

Boehm, A. E., MacLeod, D. I. A., & Bosten, J. M. (2014). Compensation for red-green contrast loss in anomalous trichromats. *Journal of vision*, *14*(13), 19-19.

Bosten, J. M., Robinson, J. D., Jordan, G., & Mollon, J. D. (2005). Multidimensional scaling reveals a color dimension unique to color-deficient observers. *Current biology*, *15*(23), R950.

Bowmaker, J. K., & Dartnall, H. (1980). Visual pigments of rods and cones in a human retina. *The Journal of physiology*, *298*(1), 501-511.

Brainard, D. H., & Wandell, B. A. (1992). Asymmetric color matching: how color appearance depends on the illuminant. *JOSA A*, *9*(9), 1433-1448.

Brainard, D. H., Kraft, J. M., & Longere, P. (2003). Color constancy: Developing empirical tests of computational models. *Colour perception: Mind and the physical world*, 307-334.

Bramwell, D. I., & Hurlbert, A. C. (1996). Measurements of colour constancy by using a forced-choice matching technique. *Perception*, *25*(2), 229-241.

CIE (1932). Commission internationale de l'eclairage proceedings, 1931. *Cambridge University, Cambridge*.

CIE (2020). *International Lighting Vocabulary.* CIE S 017/E:2020. Retrieved from http://cie.co.at/publications/ilv-international-lighting-vocabulary-2nd-edition-0

Conway, B. R. (2001). Spatial structure of cone inputs to color cells in alert macaque primary visual cortex (V-1). *Journal of Neuroscience*, *21*(8), 2768-2783.

Crone, R. A. (1968). Incidence of known and unknown colour vision defects. *Ophthalmologica*, *155*(1), 37-55.

Dain, S. J. (2004). Clinical colour vision tests. *Clinical and Experimental Optometry*, *87*(4-5), 276-293.

Daugirdiene, A., Kulikowski, J. J., Murray, I. J., & Kelly, J. M. F. (2016). Test illuminant location with respect to the Planckian locus affects chromaticity shifts of real Munsell chips. *JOSA A*, *33*(3), A77-A84.

Deeb, S. S. (2004). Molecular genetics of color-vision deficiencies. *Visual neuroscience*, *21*(3), 191-196.

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight?. *Journal of Vision*, *4*(2), 1-1.

DeMarco, P., Pokorny, J., & Smith, V. C. (1992). Full-spectrum cone sensitivity functions for X-chromosome-linked anomalous trichromats. *JOSA A*, *9*(9), 1465-1476.

Farnsworth, D. (1943). The Farnsworth-Munsell 100-hue and dichotomous tests for color vision. *JOSA*, *33*(10), 568-578.

Ferreira, Mafalda. (2017). The Information Visualization Thematic, Head to Head with Color Blindness.

Foster, D. H. (2003). Does colour constancy exist?. *Trends in cognitive sciences*, *7*(10), 439-443.

Foster, D. H. (2011). Color constancy. *Vision research*, *51*(7), 674-700.

Foster, D. H., Amano, K., & Nascimento, S. M. (2006). Color constancy in natural scenes explained by global image statistics. *Visual neuroscience*, *23*(3-4), 341-349.

Foster, D. H., Craven, B. J., & Sale, E. R. (1992). Immediate colour constancy. *Ophthalmic and Physiological Optics*, *12*(2), 157-160.

Gegenfurtner, K. R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience*, *4*(7), 563-572.

Gegenfurtner, K. R., & Kiper, D. C. (2003). Color vision. *Annual review of neuroscience*, *26*(1), 181-206.

Granzier, J. J., & Gegenfurtner, K. R. (2012). Effects of memory colour on colour constancy for unknown coloured objects. *i-Perception*, *3*(3), 190-215.

Hedrich, M., Bloj, M., & Ruppertsberg, A. I. (2009). Color constancy improves for real 3D objects. *Journal of Vision*, *9*(4), 16-16.

Helson, H. (1938). Fundamental problems in color vision. I. The principle governing changes in hue, saturation, and lightness of non-selective samples in chromatic illumination. *Journal of Experimental Psychology*, *23*(5), 439.

Hurlbert, A. (2019). Challenges to color constancy in a contemporary light. *Current Opinion in Behavioral Sciences*, *30*, 186-193.

Hurvich, L. M., & Jameson, D. (1966). *The perception of brightness and darkness*.

Ishihara, S. (1972). *Tests for colour-blindness*. Tokyo, Japan: Kanehara Shuppan.

Johnson, E. N., Hawken, M. J., & Shapley, R. (2008). The orientation selectivity of color-responsive neurons in macaque V1. *Journal of Neuroscience*, *28*(32), 8096-8106.

Kefalov, V., Fu, Y., Marsh-Armstrong, N., & Yau, K. W. (2003). Role of visual pigment properties in rod and cone phototransduction. *Nature*, *425*(6957), 526-531.

Krastel, H. (2015). *Anomaloscope - User's Manual* [Ebook] (1st ed.). OCULUS Optikgeräte GmbH, Wetzlar, Germany. Retrieved 25 September 2020.

Kulikowski, J. J., Daugirdiene, A., Panorgias, A., Stanikunas, R., Vaitkevicius, H., & Murray, I. J. (2012). Systematic violations of von Kries rule reveal its limitations for explaining color and lightness constancy. *JOSA A*, *29*(2), A275-A289.

Lagerlöf, O. (1991). The Farnsworth-Munsell 100-Hue test in clinical practice. In *Colour Vision Deficiencies X* (pp. 241-249). Springer, Dordrecht.

Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Josa*, *61*(1), 1-11.

Land, E. H., Hubel, D. H., Livingstone, M. S., Perry, S. H., & Burns, M. M. (1983). Colour-generating interactions across the corpus callosum. *Nature*, *303*(5918), 616.

Ma, R., Kawamoto, K. I., & Shinomori, K. (2016). Color constancy of color-deficient observers under illuminations defined by individual color discrimination ellipsoids. *JOSA A*, *33*(3), A283-A299.

Maloney, L. (2003). Surface colour perception and its environments [commentary]. In *Colour: Connecting the mind to the physical world* (pp. 329-330). Oxford University Press.

Maxwell, J. C. (1857). 1. Experiments on Colour as perceived by the Eye, with Remarks on Colour-Blindness. *Proceedings of the Royal Society of Edinburgh*, 3, 299-301.

Mollon, J. D., Estévez, O., & Cavonius, C. R. (1990). The two subsystems of colour vision and their roles in wavelength discrimination. *Vision: coding and efficiency*, 119-131.

Montag, E. D. (1994). Surface color naming in dichromats. *Vision Research*, *34*(16), 2137-2151.

Morgan, M. J., Adam, A., & Mollon, J. D. (1992). Dichromats detect colour-camouflaged objects that are not detected by trichromats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *248*(1323), 291-295.

Morgan, M. J., Watamaniuk, S. N. J., & McKee, S. P. (2000). The use of an implicit standard for measuring discrimination thresholds. *Vision research*, *40*(17), 2341-2349.

Morland, A. B., MacDonald, J. H., & Middleton, K. F. (1997). Colour constancy in acquired and congenital colour vision deficiencies. In *John Dalton's Colour Vision Legacy:*

Selected Proceedings of the International Conference (pp. 463-468). Taylor & Francis.

Morovič, J., & Morovič, P. (2005, January). Can Highly Chromatic Stimuli Have A Low Color Inconstancy Index?. In *Color and Imaging Conference* (Vol. 2005, No. 1, pp. 321-325). Society for Imaging Science and Technology.

Munsell, A. H. (1912). A pigment color system and notation. *The American Journal of Psychology*, *23*(2), 236-244.

Munsell, A. H. (1976). *Munsell Book of Color: Glossy Finish Collection, Removable Samples in Two Binders*. Munsell Color.

Nagel, W. (1907). Zwei Apparate für die augenärtzliche Funktionspriifung: Adaptometer und kleines Spektralphotometer (Ano-maloskop), Zsch. f. *Augenheilk*, *17*, 116-118.

Nascimento, S. M., De Almeida, V. M., Fiadeiro, P. T., & Foster, D. H. (2004). Minimum-variance cone-excitation ratios and the limits of relational color constancy. *Visual Neuroscience*, *21*(3), 337.

National Research Council. (1981). Procedures for Testing Color Vision: Report of Working Group 41.

Neitz, J., & Neitz, M. (2011). The genetics of normal and defective color vision. *Vision research*, *51*(7), 633-651.

Norman, L. J., Akins, K., Heywood, C. A., & Kentridge, R. W. (2014). Color constancy for an unseen surface. *Current Biology*, *24*(23), 2822-2826.

Oxford University Press. (2014). *Brunswik Ratio*. Oxford Reference. Retrieved September 26, 2020, from https://www.oxfordreference.com/view/10.1093/oi/authority.20110803095532286

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G. D., & Hurlbert, A. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *PloS one*, *9*(2), e87989.

Pokorny, J., & Smith, V. C. (1977). Evaluation of single-pigment shift model of anomalous trichromacy. *JOSA*, *67*(9), 1196-1209.

Pugh Jr, E. N., & Mollon, J. D. (1979). A theory of the Pi 1and Pi 3 color mechanisms of Stiles. *Vision research*, *19*(3), 293-312.

Radonjić, A., & Brainard, D. H. (2016). The nature of instructional effects in color constancy. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 847.

Radonjić, A., Cottaris, N. P., & Brainard, D. H. (2015). Color constancy in a naturalistic, goal-directed task. *Journal of Vision*, *15*(13), 3-3.

Rayleigh, L. (1881). Experiments on colour. *Nature*, *17*, 64-66.

Reeves, A. (2018). The dependence of color constancy and brightness constancy on saturation. *Color Research & Application*, *43*(5), 630-632.

Rigden, C. (1999). 'The Eye of the Beholder' - Designing for Colour-Blind Users. *British Telecommunications Engineering*, *17*, 291-295.

Rinner, O., & Gegenfurtner, K. R. (2000). Time course of chromatic adaptation for color appearance and discrimination. *Vision research*, *40*(14), 1813-1826.

Ruderman, D. L., Cronin, T. W., & Chiao, C. C. (1998). Statistics of cone responses to natural images: implications for visual coding. *JOSA A*, *15*(8), 2036-2045.

Rüttiger, L., Braun, D. I., Gegenfurtner, K. R., Petersen, D., Schönle, P., & Sharpe, L. T. (1999). Selective color constancy deficits after circumscribed unilateral brain lesions. *Journal of Neuroscience*, *19*(8), 3094-3106.

Rüttiger, L., Mayser, H., Sérey, L., & Sharpe, L. T. (2001). The color constancy of the red-green color blind. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, *26*(S1), S209-S213.

Saito, A., Mikami, A., Hosokawa, T., & Hasegawa, T. (2006). Advantage of dichromats over trichromats in discrimination of color-camouflaged stimuli in humans. *Perceptual and motor skills*, *102*(1), 3-12.

Shrestha, R. (2016). Simulating colour vision deficiency from a spectral image. *Stud. Health Technol. Inf.*, *229*, 392-401.

Simunovic, M. P. (2010). Colour vision deficiency. *Eye*, *24*(5), 747-755.

Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1329-1346.

Tyler, J. E., & Hardy, A. C. (1940). An analysis of the original Munsell color system. *JOSA*, *30*(12), 587-590.

Von Kries, J. (1878). Beitrag zur Physiologie der Gesichtsempndungen [Transl. Physiology of visual sensations]. *Sources of Color Science*, 101-108.

Wan, L., & Shinomori, K. (2018). Possible influences on color constancy by motion of color targets and by attention-controlled gaze. *JOSA A*, *35*(4), B309-B323.

Wandell, B. A., Baseler, H., Poirson, A. B., Boynton, G. M., & Engel, S. (1999). Computational neuroimaging: color tuning in two human cortical areas measured using fMRI. *Color vision: from genes to perception*, 269-282.

Weiss, D., Witzel, C., & Gegenfurtner, K. (2017). Determinants of colour constancy and the blue bias. *i-Perception*, *8*(6), 2041669517739635.

Werner, A. (2014). Spatial and temporal aspects of chromatic adaptation and their functional significance for colour constancy. *Vision research*, *104*, 80-89.

Westheimer, G. (1965). Spatial interaction in the human retina during scotopic vision. *The Journal of physiology*, *181*(4), 881.

Willmer, E. N. (1961). Human colour vision and the perception of blue. *Journal of Theoretical Biology*, *1*(2), 141-179.

Worthey, J. A. (1985). Limitations of color constancy. *JOSA A*, *2*(7), 1014-1026.

Worthey, J. A., & Brill, M. H. (1986). Heuristic analysis of von Kries color constancy. *JOSA A*, *3*(10), 1708-1712.

Xiao, B., Hurst, B., MacIntyre, L., & Brainard, D. H. (2012). The color constancy of three-dimensional objects. *Journal of Vision*, *12*(4), 6-6.

Zaidi, Q., & Bostic, M. (2008). Color strategies for object identification. *Vision research*, *48*(26), 2673-2681.

Zeki, S., & Marini, L. (1998). Three cortical stages of colour processing in the human brain. *Brain: a journal of neurology*, *121*(9), 1669-1685.